

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTTCAGCCCTTGTAAATTGGACATCTGCTGCTTCAATT
TCATACATTACTGCAGTAACACTCCACCATAAGACCCGGCTTACCTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTATGCATTG
CTACCATTTATGTCGTTAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAAGGCTGGCCTTGTACTTGAATACTGAGTTAGGACTTCTATTGTCGAAACTT
CCAGAAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTCAGACCATCCTTCCTACCAAATGCAGCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTGTCACAGTGGCAATTGGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGCTATGTCATTCCCTT
GGTTTTCTGACTTACATTGCTGATTTCAAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCTCTATGACACTGCACCTGCCCTATTACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGGGTGGCGGACCGTGGGGAGAGCCGAGTCGGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGGAAGTGGAGGCAGGCCTTC
CTTACACTCGCCATGAGTTCCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTCTTCATGCGCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGACAG
GTGATCTCTCCGTGACGTTGCATTTCTGCACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTTCACTGGAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTCATGGTGCCTTTTACATTGGCTATTTATTGTGAGCAATATCGACTACTGCATAAACAGA
CTGCTTTCTGTCTTATGGCTGACCTTATGTATTCTCTGAAACTAGGAGATCCCTTCC
CATTCTAGCCAAAACATGGATCTTATCCATAGAACAGCTCATCAGCGGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGTGTCACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGAAAAAGAAAAGGATGCAATGGCACGGAGAACATGTCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGAAATATTTAATTCTGGTT
ACTTTTCTCTATTACTGTGTTGGAAAATTTCATGGTACCATCATGAGTT
GGGAAAACGGATCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTTCATTCTGTTGAAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTCTTTATGCATCTCTAGCAACTGCTCCTGATCCGAATGAGTATGCCTT
CTATTAGCACAGATAATGGCATGTACTTGTCTCCTGTGCTGATCCGAATGAGTATGCCTT
AGAATACCGCACCATAACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAA
GATATAAGAGGGGGAAAATGAAACCAGGGCTGACATTATAACAAACAAAATGCTATGGTAGC
ATTTCACCTTCATAGCATACTCTCCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCAAGAAACTAAAGGTGAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTTAGGATTTCCGTTAAGGTTACATGGAAAGGTTATAGCTTGC
CTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCCATGGCCAACCTGTTATTGCAGCTATAATG

FIGURE 4

MSFLIDSSIMITSQLILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGDP
FPILSPKHGILSIEQLISRVGVIGVTIMALLSGFGAVNCPTYMSYFLRNVTDILALERRLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSCFWGMIKSVTTSASGSENLTLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGTKDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQMIMGMY
FVSSVLLIRMSMPLLEYRTIITEVLGELOQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCATCCGACTAGTTATTGAGCATCTGCCCTCATATCACCAAGTGGCCATCTGAGGT
GTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCAAGGCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCCTGTGAGCAAAAAGGCGAACCGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCGTAGGCTGCTGGGACTAAGTTGGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTGTGTCATCTCTAG
GATTAGCCCCAACCCCAAGGTGGGAAAATGGGGTGGGTGTCTGATTGGAAAGGTTCCAGTGA
GCCGACAGTTGCGCTATTGTTACAACACTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
ATTATCACCAAAAGATCCCATTCAACACTCAAACACTGCAACACAAACACAGAATTATGT
CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCTACTACTCCTC
CTGCTCCAGCTTCCACTTCTATCCACGGAGAAAAAAATTGATTGTGTCACAGAAGTTTATG
GAAACTAGCACCAGTCTACAGAAACTGAACCATTGTGAAAATAAGCAGCATTCAAGAATGA
AGCTGCTGGTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCCTTGGTGTG
CAGCTGGCTTGGATTGCTATGTCAAAAGGTATGTGAAGGCCCTCCCTTACAAACAAGAAT
CAGCAGAAGGAAATGATGAAACCAAAGTAGTAAAGGAGGAGAAGGCCATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACCTACCGTGC
GATGCTGGAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCTTACCCGCCAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAACAGAAAGTCCA
CCCTGGTTCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
GCCCTTCTCTTATGTAAACCTGTCTGGATCCTATCTCTACCTCAAAGCTCCCACGGCT
TTCTAGCCTGGCTATGTCTTAATAATATCCACTGGAGAAAGGAGTTTGCAAAGTGAAGGAC
CTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTGGCTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCTCTGAGCCCGTAACAGCAAAAGAAT
GGCAGAAAAGTTAGCCCCCTGAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAAC
GCTAAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGCTCTACCACTGCTGATATTTCTCTAGGAAATAACTTTACAAGTAACA
AAAATAAAACTCTTATAAAATTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAATATAAAATCAACAAACATTGAGATTGCTGAATAGCTACTATATGTC
AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATTGAAATTATTCCCTAAAAAATTGCA
AACATAGTAACTTTCTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTCTGCTGAGACTAATCTTATCTGATATTCTCTAATATGCAACATTATAACCTTAATT
TATTATTAACATACTAAGAAGTACATTGTTACCTCTATACCAAGCACATTAAAAGTGC
ATTAACAAATGTACTAGCCCTCTTTCCAACAAAGAAGGGACTGAGAGATGCAGAAATATT
TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYCWVGDFVVISRISPNCGKNGVGVLIWKPVRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAALKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCCTCCGCACCCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCCGC
GGCCTCCC GGCGGGAGCGAGCAGATCCAGTCGGGCCGCAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTGCGGGCGCAGAGCGGAGATGCAGCGCTTGGGCCACCCCTGCTGTGCCTGCTGCTGGCG
CGGCGGTCCCCACGCCCGCGCCGCTCCGACGGCAGCTCGGCTCCAGTCAGCAAGCCCGCCCG
GCTCTCAGCTACCCGAGGAGGCCACCTCAATGAGATGTTCCGCAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGACCGAGAAATTCAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTCAGAGACAGTTACATCAGTGGGAGACAAGAAGGCAGAAGGAGCC
ACGAGTGCACTCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACCAAATGCCACAGGGGAGCAATGGGACCATCTGTGACA
ACCAAGGGACTGCCAGCCGGGCTGTGCTGTGCCCTCAGAGAGGCTGCTGTCCCTGTGTC
ACACCCCTGCCGTGGAGGGCAGCTTGCCATGACCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGGCCCTGATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGCCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCCTGCCAG
CCGCCGCTGCACTGCTGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGTAGATGTGAA
TAGAAATAGCTAATTATTCCCAAGGTGTGCTTAGGCGTGGCTGACCAGGCTCTCTA
CATCTCTCCCAGTAAGTTCCCCTCTGCCCTGACAGCATGAGGTGTGCAATTGTCAGCT
CCCCCAGGCTGTTCCAGGCTCACAGTCTGGCTTGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATGGCTGCTTGCCCTACCAAGTTGGCAGACAGCG
TTTGTCTACATGCCCTTGATAATTGTTGAGGGAGGAGATGGAACAAATGTGGAGTCTCCCT
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAACACATCAACCTGGCAAAATG
CAACAAATGAATTTCACGCAGTTCTTCATGGGCAAGGTAAGCTGTGCCTCAGCTGTGC
AGATGAAATGTTCTGTTCACCTGCATTACATGTGTTATTCACTCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCAAGATCAATTCCCTCTCAGCACAGCCTGGGG
AGGGGTCATTGTTCTCTGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTTGTGACTCTAACGTCAGTGTCT
CTCCACTACCCCACACCAGCCTGGTGCACCAAAAGTGTCTCCCAAAAGGAAGGAGATGGGAT
TTTCTGAGGCATGCACATCTGAATTAGGTCAAACTATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGGCTCAGAGACTGCAAGCTGCTGCC
TACAGGTTAACCTGCAGAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACAGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTGTAAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTCACTGGACTGTTGCCACCATGTATTCACTCCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATGTATAAGCATGTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAASSEVNLANLPPSYHNETNTDKVGNNTIHVREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMCLTRDSECCGDQLCVWGHC
TKMATRGSGNTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGIELLPREVDEYEVGSFMEEVROELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGCCAGGGCTGATGTGCGTCTTCAGGGCTACTCATCCAAAG
GCCTAACCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCTGGGCTCTCTGGACCCCTT
AACTGGGTACTGGCCCTGGCCAATGCGCTCGCTGGAGGCCCTTGCCCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCCCTTAATCTGCCTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGGCCCTACCTGCCTGACCCCTGTGCAGATAGCCCAGGTACATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCCTGCATCATGTGCTGTTT
CAAGTGCCTGCCTCTGGTGTCTGGAAAAATTATCAAGTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTGTCTAGCCAAAATGCGTTCATGCTACTCATGCGAAC
ATTGTCAAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTGGAAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGCTCTTTCTCCGGTCGCATCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCTGGGGCCTAT
GTCATGCCAGCGGCTCTCAGCCTTCAGCGTTTCCGGCATGTGTGTCAGCCTCTGCTGCTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGAAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTGTGGTAAAAAAAGTTTAGGCCAGGCGCCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCCGGGCGATCACCTGAGTCAGGAGTTCGAGGACAGCCTGCCAACATGGTG
AAACCTCCGTCTTAAACAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGAGGCTGAGGCAGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTTATTAAAGATATTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLFFFGLKLLVVGVGVLSSFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCGGCGCCGGCGCCCCGAAGCCCCGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCCGTCTGCCCTCTGCCGTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTCACGTTCCCTCTTCCCTGGGTGCTGGTCCA
TCATTATGCTGAGCCGGCGTGAGAGTCAGCTACAAGCTGCCCTGGGTGAGGAGGGGCC
GGGATCCCCACCGTCTGCAGGGCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCCTCTCTCTTTAACCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGTTTGGTCTTAAGTCCTGATCCTGGTGGCCTC
ACCGTGGGTGCCCTACATCCCTGACGGCTCCCTCACCAACATCTGGTCTACTTCGGCGTGG
CTCCTCCCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCCTTCTTCACTCTCCTC
TTCTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTCATGTAACACTGAGCCCAGCGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGTCGCGTGTCCATCGTGTGTCTGC
CCAAGGCCAGGACCCCCAGCCCCAACCTGGGCTGCTGCAGGCCCTCGGTACACCCCTACACCATG
TTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCCCTCATCATCTCCCTCTGTGCACCCCTTCATCAGTCTGGCTCCTCAGACACCAGG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAG
GGCAGCCTGTGAGGGCCGGCCTTGAACAACGAGCAGGACGGCGTCACCTACAGTACTCCTTCTC
ACTTCTGCCCTGGTGTGGCCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGCCCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCCAACCGCAGTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGCCTCTGCCACCTGGTGCCTCGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCCACCCCTGCCCAAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCCTTCCTCCCTCCCTGGTGTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTGTCTCAGGCTCACGGGAGCGGGCTGCTGGAGAGAGCGGGGAACTCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTACGTCCCCCAGGGACCCCTGCCCTTCTG
GACTTCGTGCCCTACTGAGTCTCTAAGACTTTCTAATAAACAAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCS CCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHAWNQRWLKGAE
ECDSRAWYAGLFFF TLLFYLLSIAAVALMF MYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSG LQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETV VAGPEGYETQWW DAPSI
VGLIIFLLC TLFISLRSSDHRQVN SLMQTEECPPML DATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLV LASLHVMMT LTNWYKPG ETRKMISTWTAVWVKICASWAGLLLWTLVAPLLRN RD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTCTGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCAGTCTCTCTGCAGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTCTGTTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTCTCTGGCAGTTTCGATTTAAAGTGTAAACTTGCATATGCTGTGTCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAG
TGATCCTTCGAAGCTTCTCAAGGGGTTTGGCTATGTGTCGCCATCATTCATTCA
CTTGCCTGGATTGAGACGTTGGCTGGATTCAAAGTGTACCTCAAGAACAGAGAACAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTATTCCCCCTCCTGAATCGAACAGCAGGATCTGAAGAACAGCTGAAGAAAACAGGACAGT
GAGAAACCACTTTAGAACTATGAGTACTACTTTGTAAATGTGAAAACCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCCCTGTCGACAGTAAAGTGTGAAATGGTACGTC
CACTGCTGGCTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATACTTTTATTCAATTAC
AGAATGGAATTTTGTTCATGTCAGATTATTGTATTCTTTAACACTCTACATT
TCCCTGTTTAACTCATGCACATGTGCTTGTACAGTTAAAAGTGTAAATAAAATCTG
ACATGTCATGTGGTAGTTATTCTGTGTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTTCTGTGATTGTTCTGAAATGTTCTAAATACTCTTATTGTAATG
ACAAAATGACTAAACCATTCAATCATGTTCCCTTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWAIALT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTCTGGGACCCAGGACCCCCCTGGGCCGACCCGCCAGGAAAGACTGAGG
CCGC GG CCTGCCCGCCGGCTCCCTGCGCGCCGCCCTCCCGGACAGAAGATGTGCTCAG
GGTCCCCTCTGCTGCGCTGCTCTGCTACTGGCCCTGGGCCCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAAGCCACAGACAGTCTCTGCACTGCCGCCAGGGGACCACGGTGCC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATACCATGCTGACGC
AGGCAGCTTGGCCCTGGCGCTGCGACCTCTGGACCTGTCACAGAACCAAGATGCCAGCC
TGCCCAGGGGCTTCCAGCCAACCTCGCCAACCTCAGCAACCTGGACCTGAGGCCAACAGGCTG
CATGAAATACCAATGAGACCTCCGTGGCCTGCGCGCCCTCGACACGCTCGACGCCCTCTGGAC
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACGCCCTCTGGAC
AGGACAACGAGCTGCCGGCACTGCCCCCGCTGGCCTGCCGCCCTGCTGCTGGACCTCAGC
CACAAACAGCCTCTGGCCCTGGAGCCCGCATCCTGGACACTGCCAACGTGGAGGCCTGCGGCT
GGCTGGTCTGGGGCTGCGCAGCTGGACGAGGGCTCTCAGCCGCTTGCACACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCCTGGGGGCTGACG
CGCCTGCGGCTGGCGGCAACACCCGATTGCCAGCTGCGGCCGAGGACCTGCCGCCCTGGC
TGCCCTGAGGAGCTGGATGTGAGCAACCTAAGCCTGAGGCCCTGCCCTGGCAGCCTCTGGGCC
TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCGCAACCCCTCAACTGCGTGTGCCCCCTGAGC
TGGTTGGCCCTGGGTGCGCAGAGCCACGTCACACTGGCCAGGCCCTGAGGAGACGCCGCTGCCA
CTTCCCGCCAAGAACGCTGGCGGCTGCTCTGGAGCTTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCACCCAGGAGGCCCTGGTGCAGGCCACAGCCTTGTCT
TCTAGCTTGGCTCTACCTGGCTTAGCCCCACAGGCCGGCACTGAGGCCAGGCCCT
CACTGCCCAACCGACTGTAGGGCTGTCCTCCAGGCCACTGCCACCGTCACCTGCCCTCA
ATGGGGCACATGCCACCTGGGACACGGCACCCATGGCGTGTGCTTGCCCCGAAGGCTTCAG
GGCCTGACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGGCCCTACACCAGTCAGCCAG
GCCACCACGGTCCCTGACCCCTGGGATCGAGCCGGTGAAGGCCACCTCCCTGCGGTGGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCACTCAGGAGCCTCCGTCTCACCTATCGAACCTATCG
GGCCCTGATAAGCGGCTGGTACGCTGCACTGCCCTGCCCTGCCCTGAGTACACGGTCACCCA
GCTGCGCCAACGCCACTTACTCCGTCTGTGTCATGCCCTGGGCCCGGGGGCTGCCGAGG
GCGAGGAGGCCCTGGGGAGGCCATACACCCCAAGCCGCTCCACTCCAACCCAGGCCAGTCACC
CAGGCCCGAGGGCAACCTGCCGCTCTCATGCCCGCCCTGGCCGCGGTGCTCTGGCCG
GCTGGCTGCCGTGGGGCAGCTACTGTGTCAGGCCGGGGCCATGGCACCAAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGGCTGGGCCCTGGAACTGGAGGGAGTGAACGCTCCCTGGAG
CCAGGCCGAAGGCCACAGAGGGCGGTGGAGAGGCCCTGCCAGGGCTGAGTGTGAGGTG
ACTCATGGCTTCCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAGGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGGGGCTCTCAGCCAGTGAGATGGCCAGGCCAGGGCTGCTG
ACACCCAGTAAGTTCTCAGTCCCAACCTCGGGATGTGTCAGACAGGGCTGAGTACACG
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCATCTGAGATGCTGTCAGGCCAGCTGACGAGCC
CTAACGCTCCCAAGAACCGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCAACGTGCA
CCTGGGCACGGCGGGCCCTGCCATGTGTCAGGCAACGGCTGCTGGCTCTGGCTCTCCCAC
TCCAGGGGACCCCTGGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTTGGCCCCAGGAAGCGAAGGAACAAAAGAAAAGTGGAAAGGAGATGC
TTTAGGAACATGTTTGTCTTTTAAATATATATTATAAGAGATCCTTCCCATTTATTCT
GGGAAGATGTTTCAAACCTCAGAGACAAGGACTTGGTTTGTAGACAAACGATGATATGAA
GGCCTTTGTAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDSLQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLDNELRALPPLRPRLLLLDLHSNSLLALEPGILDGTANVE
ALRLAGLGLQQLDEGLFSRLRNLDLSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLALPGDLSGLFPRLLAARNPFCNCVCPPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSLRVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASIAEY
TWTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAAYCVRGRAMAAAQDKGQVPGAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGCGGTGGCTGAGTCGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGCGGTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTAGAAGAAATCTGA
ATTAGAACCTCTATTCAAGAACAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAAGTA
CGGAAACCAGCTTGACCGCCATTGAAGGCACAGCACATGGGAGGCCCTGCCACTTCCCTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGT
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAACAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAACAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCAATGCTCTTATTGGTGAATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTCTGTATATTACAT
TTGGAGCTTGGGGCAATCTAAAGCCCACATGGTTTGGTAAGTAGACTTTAGTGGAAAGCT
AATAATATTAACATCAGAAGAATTGTGGTTATAGCGGCCACAACCTTTCAGCTTCAATGATC
CAGATTGCTGTATTAAGACCAAATATTCAAGTGAACCTCCTCAAATTCTGTAAATGGATAT
AACACATGGAATCTACATGTAATGAAAAGTGGGGAGTCCACAATTTCCTTAAATGATTAG
TTTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAAATGGCTCTTTAAATTCTCTGAGTTG
GAATTGTCAAGATCATTTCACATTAGATTATCATAATTAAATTTCTTGTAAATGGATAT
AAATTGTAAATGGGGCTATAGAAAAACACATGAAATATTACAAATTGGCAACAATGC
CCTAAGAATTGTTAAATTCACTGGAGTTATTGTGCAGAATGACTCCAGAGAGCTCTACTTCTG
TTTTTACTTTCACTGATTGGCTGTCTCCCATTTATTCTGGTCAATTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTCCTATTGTCAATTGTACTTTCTTGTAAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTTTTTTTTTTTTTTTTTTTTTTTTTT
AA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTTLTSDESVKDHTTAGRVVAGQIFLDSEEESEL
ESSIQEEEDSLKSQEGERVTEIDFLESPNPNENKDYEPPKKVRKPALTAIEGTAHGEPCFPFLFLDK
EYDECTS DGDREDGRWCATTYDYKADEKWGFCETEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAAGATTTAACGCCATTCTGCAGTGGAAATTCACTGAACTAGCAAGAGGACACCATTCTT
GTATTATAACAAGAAGGAGTGTACCTATCACACACAGGGGGAAAATGCTCTTGTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTCTGTGGACTCGTAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATTCATGTAATCGCTGCCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCAGCCTGGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCCTCAGTGTGAGGTCGCCTTGCATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAAATTGAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAACATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCAACAGCTCTCCCTAACACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTGAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAACCCAAAGGCAGTTGACTCAGCTAACACAAATGTCCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCCTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGTCTATTGGATTGCAAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTAGGCTTGCCTGCTTGGTGTGATGTAAGGAAATTGAAAGACTTGCCATTCAAAATGA
TCTTACCGTGGCCTGCCCATGCTTATGGCCCCAGCATTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGKGLWGLINNAGVPGVLAPTDWLTLLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCLIEPGLFKTNLADPVKVIKEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGTTTCCAGCACCTAACACCGACTCGGACACCGAAGGTTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACATTGA
CATTCAAGAAAATATATTCCATGCTATCAGCTTTAGCTTATAATTCTTCAGGCGAAGTAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAACTTGCAGGAGCATT
TTCAAAACCAAGACCTTGTTCTGCTATTAAACACCAAGTATAAAACAGAAAGCTGCTACTC
ATCGACTGGAACATCCTTATATAAACCTCAAAAAGGACTTTTCACAGGGTACCTTAGTGGTT
GCCAATCTGGCATGTCTGAACAACGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACAGCTCTAAATTGGAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTTCTTGTCAAGCATTACGGACCTTTTCAAATTCTGAATTCTTCAATTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACTACAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTGGTAATATTACCGTCTC
CTACATTTGATCCTTAACCTTACAAGGAGATTTTATTGGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTGAGCTACTTGCAGTAAGTCATTGTTTACTATGTTCA
TGTTGAGTAATACACAGATAACTCTTAGTGCATTACTTCACAAAGTACTTTCAACATCA
GATGCTTTATTCCAAACCTTTTCACTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTACAAGTGGAAAAGTGAGACCAGGACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGCTATTAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAATTATCTGAGTCATTAAATTCTCCTTAAGTGATAACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTSGSCMSTGFSRAV
QTHSSKKFEEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVKSSCNYNHLDVVDNLTL
MVEHTDIPPEASPASTPQIIKHKALDDLDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGACGCCGAGCGGACCAGCCAGGGCAGCCCAA
GCAGCGCGCAGCGAACCCCCGCGCCGCCACACCCCTCGCGTCCCCGCGGCCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCGCGCTCGCGGCCAGTCAGCTTGCGGGTCTCGCTGCCCGCGAAACCCCGAGGTCAACAGCCCGCGCTCT
GCTTCCCTGGGCCGCGCGCCCTCACGCCCTCCTCTCCCGCCCTGGCCCGGCGCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCATTTGCCCGCTCTCCCGCTCGCTGCCCTCTCCACCAACTCCAACCTCCCTCTCCC
TCCAGCTCCACTCGTAGTCCCCGACTCCGCCAGCCCTCGGCCGCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCAAA
GGTGGGAACCGCTCCGCCCGGCCGACCAATGGCAGGTTCGGCTTGCCCGCGCTCTGCACCCGGCAGTGC
AGCGCCGCGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCAGCTTACGTGTCAAAGGCTTC
AACAGAAGCATGCCCTCCACGAGATCAACGGTATCAATTGAAGATCTGCCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGAATCATTG
CAAGCTGTCTTGCTTCAGTTACAAGAAGTTGATGAATTCTCAAAGAAACTACTTGAAAATGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGCCATTATACATGCAAATCTGAGCTATTAAAGATCTCTCGTAGAGTTG
AACGTTACTACGTGGTGGAAATGTGAACCTGGAGAAAATGCTAAATGACTCTGGGCTGCCCTGGAGCGGATG
TTCCGCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGAATGTTGAGCAAGTATACGGAGCAGCTGAAG
CCCTCGAGATGTCCCTCGAAATTGAAGCTCCAGGTTACTCGCTTTGTACCGAGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGAECTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTGGCCAACCAAGGGATCTGATTTGAATGAAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCCTTCACATTGAATCGGTATGGATCCCATCGATGTGAAGATTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTCAGTGTCTCGCTTCAGACCACATACCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTTACTGATGTCAGGAGAACTGAAACAGGCCAAGAAATTCTGGCCTCCCGAGCAAC
GTTTGCACCGATGAGAGGATGGCTGCAGGAAACGCCAATGAGGATGACTGTTGAATGGAAAGGCAAAGCAGGTAC
CTGTTGCAGTACAGGAAATGGATTAGCAACCAGGGCAACACCCAGGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCTCGTCAAATCATGGCTCTCGAGTGTGACCAAGATGAAGATGCTACAAATGGAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTCCCTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAGAGTGCCTGCAATGAGAAAGCCGACAGTGTGGTCCGTCTGGGCA
CAGGCCTACCTCCTCACTGTCCTGCACTTGTCTGGTTATGCAAGAGAGACTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTCAAAAGTTAAAGGCACCGTTATCACTTTCTACCATCTAGTGACTTGCTTTAAATGAA
TGGACAACAATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGTGACTTGTTCTCATTCAGTTGG
AGGAAAGGGACTGTGATTGAGTTGGCTCTGCTCCCCAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTGCTTACACTCTATTATTTGTTATGTTTTCTCATTCGTTGTGGGTT
TTTTTCCAACTGTGATCTCGCTTGTCTTACAAGCAAACCAAGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAATAGCTGTACAGAAGCAGGTTTATTTATCATGTTATCTTAAAGAAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKCSEVRRLYVSKGFKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNLQAVFASRYKKFDEFFKELLENAEKSNDMFVKTYGH
LYMQNSELFKDLFVELKRYYVGVNLEEMLNDFWARLLERMFRILVNSQYHFTDEYLECVSKYTE
QLKPGDVPRLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVNPTAQCTHALLKMIYCSHCRL
VTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNVNCNDERMAAGNGNECDCWNGKGSRYLFATVGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVFFFISDESSGEKGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glycans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGCGCCAAGAATGTGAGTGAAAGATTGGTCCTGAG
AGCCCCGAGAAGAAAATTATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAAGGCACCACAGAAAGCCAACAAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACATGTCAGCTAAGAAGCTTGCTGCCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTAACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSLNPVGARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPDCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCTGCAGCTGCTGGTGTGCTTCTTACCCCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTTCAGCCAG
ATAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCTACTGGAGCTGGCTGCGGAACCGGAGC
CAACTTCAGTTCTACCCACCAGGCTGCAGGGTACCTGCCTAGACCCAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTGTGACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTTCTGGGAGCATGTGGCAGAACCATATGAAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCGAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGCTACCTGTT
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCTATCTATCTTCCACTGAGAGGGACCTA
GCAAGATGAGAGAACATTGATGTACCCACTAGTCCCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCCTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCTC
CCAATGTTGCCCTTCCCTCGTTCCATGGTAAAGCTCTCTCGCTTCTCTGAGGCTACAC
CCATGCGTCTCTAGGAACCTGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTCCCTGAGCTGGGGGCCACCAGGGAGAATCAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTAATAATAGACGAA
ACCACG

FIGURE 28

MDILVPPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCSVQSPRKVLQEVRVRLPAGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGGCCTCTTCCCCCTCCTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGAGATAGGAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTTATTTAAA
GTGGTTTTATGATTCTTATACTAATTATAACAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSSAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCAGATTCCCTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTTATTTAACCGATGTGTCCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTGGTTCACATTTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTCCGTGCTGGGTTGCTATCGTATCCACAGGCATACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGGTCTCTGGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAAATATAAGCCCCTTCGGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCACAGAACTAAAAATGATCCTCTGATCATCCCATCCTTCGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGTGAAAGGGCATTAAATCTGTGGTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTCGATGCTGACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAGATGCATAACTACAACGTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGTCCAAGAACTCAAGTCACTTACATCTATTAACTGCTTGAGACTTATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCAGAAACCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACTTTCTGTGTTGCTGATGATC
TGGAAACAAATGATGGATCGTCAGAAACCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATGAGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCCTCTAAGAGCCATTACAGAAATAGAAGATGAGACCCTAGAGAAAAGTTAGTGAATT
TTAAAGACCTAATAAACCTATTCTCCTCAAA

FIGURE 32

MSGRTDTILGLCILALALS LAMMFTFR FITTLLVHIFISLVLG LLLFVC GVLWWLYYDYTN DLSIE
LDTERENMKCVLGFAIVSTGITA VLLVLIFVLRKRIKLTVELFQITNKAIS SAPFLLFQPLWTFA
I LIFFWVLWVAVLLSLGTAGAAQVM EGGQVEYKPLSGIRYMWSYHLIGLIWTSE FILACQQMTIA
GAVVTCYFNRSKNDPPDHPI LSSLSILFFYHQGTVVKG SFLISVVRIPRIIV M YM QNALKEQQHG
ALS RYLFRC CYCCFWCLDKYLLHLNQNAYTTAINGTDFCTS A KDAFKILSKNSHFTSINC FG D
FI IFLGKV LVVCFTVFGGLMAF NYNRAF QVWA VPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAI VR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCCTGAGAAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTTCCTAGA
ATAATTGTATGGGATTTGTGATGCAGGAAGCCTAACAGGAAAGAATATTCAATTCTGTGTGGT
GAAAATTTTGAAAAAAATTCGCTTCTTCAGAACAGGGTGTCAATTCTGATATTTATGAGGAC
TGTTGTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGACTGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAGGCCAAGTCACTGTGCTCATGACTG
GATGTCAAAGCCGAAAGATCATCGATCTGAGTCAATTGTGAAATGTCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGATGCATCCTACTCCAGTGTGTCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGAAATACTGTTCTGGAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTATTCCAAGGTGTCCAATCGTATACCTACAGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCAAAAGGGTGTAACTACCCATCAGCTTACATACTCAT
CGAAAAGTCCAGTGCCAAGCAGGTGAGACACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACAGCACAGCCGGTCACTCTGATGAGCTTCTGGCTGCACTGTAGCTGTGGCCACCCAC
CACCTTGCCAAGGCCATCCCCTCTGCTGCTTCAACCAACAGATCCCAAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCACTGCCACCTACACAAGCAGC AAAACAGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCCTCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGCTTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGCTTTTAATTGATGGAGCACCAGCATGGCAA
CGGCGATTCCGAATCCAGAAGCAGCTCTGGCTGATGTCAGCTTCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTGTCAGTATGGAGACAACCTGCTACTCACTTAAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCTTGTGACCAAGAACCTTCTTCCAAAGCCAATGAAAACAGAACGG
GGCTCCAATGTGGTGGTGGTGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTTC
GACTTGGCAGAGAGTCAGGAATCAACATTTCCTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCAACTTTGCAAACAAAGGGCTGTGAGAACAAACGGCTTCACTC
GCTCCACGTGCAGAGCTGGTTGGCCTCACAAGACCCCTGCAAGCTCTGGTGAAGCGGGCTGCG
ACACTGACCGCCTGCCCTGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTACGAC
GGCTCCAGCAGTGTGGGACGGCAACTCCGCACCGTCCCTCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTTCGACACGGACACGGCATCGGGCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGGTTGCAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGGG
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAGAGGAATTAAATGATCCTCATCACCGACGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCCCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGAACACTCCTTCTTGAGCGA
GTTTGACAACCTCCATCAGTATGTCCTCAGGATCATCCAGAACATTGATCAGAGTTCAACTCAC
AGCCTCGGAACTGATTCAAGAGCAGGGAGAACAGCAAGTGTGCTTTACTAACTGACGTGTT
GGACCAACCCACCGCTTAATGGGCACGCACGGTGCACTGAAGTCTGGGAGGCCATGGAGAAC
AAATGTTGTTATTATTCTTGCATCATGCTTTCTATATTCCAAAACATTGGAGTTACAAAGA
TGATCACAACAGTATAGAATGAGCAAAGGGTACATCATGTTGAGGGTGTGGAGATTACAT
TTTGACAAATTGTTCTAAATAATGTTGGAATACAGTGCAGCCCTACGACAGGCTTACGTAG
AGCTTGTGAGATTAAAGTGTATTCTGATTTGAACTCTGTAACCCCTAGCAAGTTCACT
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAATAAAAAAA
AAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
PTTLPRLPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADEPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQPLVVKRCVDRLACSKTCLNSADIGFVIDGSSSVGTGNFRVTLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSCKPNKRKLMILITDGRSYDDVRI PAMAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGCTTTTCTGTTCAGAGTAGTTACAACAGATCTGAGTGTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTTCATCTGAAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTCTACACGTGGTGCTCTCCGACTACTCACCCCGAGTGT
AAGAACCTCGGCTCGCGTCTGAGCTGCTGTGGATGGCTCGGCTCTGGACTGTCTCCGAGTA
GGATGTCACTGAGATCCCTAAATGGAGCCTCTGCTGTCACTCCTGAGTTCTTGTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTAACTGGATGACTTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACTTCACACTCGAGAGCATTCAAACAGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCACCCTCAGATGTGAAAGCCAGGCCATTAGAGTTACTGGGGTGA
TGGTGGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAATGTTGG
ATTGTCCTTAGAGGATGAACACCTCTTATGGTACATAATCCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGA
AAGACAGACACTGATGTTCATCAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTACAGGTATCCTCTAATTGATAATTATTCTATAGAGGATTACCAAAAAACCCATATT
CTTACCAGGAGTATCCTTCAAGGTGTTCCCTCATACTGCAGTGGGTTGGTTATATAATGTCAGAG
TTGGTGCCAAAGGATCTATGAAATGATGGTCACGTAAACCCATCAAGTTGAAGATGTTATGTCGG
CTGTTGAATTATTAAGTGAACATTCTATTCAGAACAGACACAAATCTTCTTCTATAGAATCC
ATTGGATGTCTGTCACAGACTGAGACGTGTATTGCAAGCCCAGGGTTCTTCCAGGAGATCATC
TGGCAGGTATGCTAAGGAACACCACTGCCATTTAACTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTGTGAAAGTGTAAATAAGTAGGACTGTGGAAAATTCTAGGGAGGTCA
ACACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATT
CCCTCAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAGG
ACCAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAGGGTTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAAACATGTAGAGTTTATTGAAACAATGTAGTCACCTGAAGGTTTG
TATCTTATGTGGATTACCAATTAAATATGTAGTTCTGTCAAAAACCTCTCACTGAAGTT
CTGAACAAAATTTCACCTGTGTTGGCATTATAAGTACTTCAGTCAAGATGTT
ATTATTTAAATTACTTCACCTTGTGTTAAATGTTGACGATTCAACAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACGTGATCAGTTATTGATA
TTAATGAAAGTCATAGGTCAATTGCAATATCAGTAATCTTGGACTTTGTTAAATTTACTGTGG
AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPF LVILVTSHPSDV KARQAI RV TWGEKKS WWG YEV LTF LLG QEA EKD KMLA
LSLEDEHLLYGDII RQDFLDTYNNLTLKTIMAF RWVTE FCPNAKYVMKT DTDV FINTGNLV KYLL
NL NHSEK FFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMS RDIVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHI PEDTNLFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCGCCGGCAAGGATGGAGCTGGGTGCTGGACGCAGTGGGGCTACTTTCTCAGCTCCTCTCATC
TCGCCTTCCAAAGAGAGTACACAGTCATTAATGAAGCTGCCCTGGACAGAGTGGAAATATCATGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTGCCCGAAAGAGGAAAGCTGTGGTTATACCATCCCTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCCTGCCATCCACCCAGGTTGACTCATCTTGAAAGACTGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTGTGGAAAGCTATCCCCTAAATGTCACTGTGAATGGACCATTGATGC
TAAACCTGGGTTGTCACTCAACTAACGATTGTCACTGTTGAGCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCGTGATGGAGACAACCGCGATGCCAGATCATCAAGCGTGTCTGGCAACGAGCCGAGCTCTATCCAGACATAG
GATCCTCACTCCACGTCCTTCCACTCGATGCCATCAGAATTTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCTCATCCCTGTTCCATGACGGCACGTCCTGACAAGGCTGGATCTTACAAGTGTGCCCTGCTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCTTGAAGAAAAGAAACTGTCAGACCCCTGGGGCCAGTCATGGTACAGAAAAA
TAAACAGGGGGCCCTGGGTTATCACGGACGCCATGCTAAATTGGCACCGTGTGTCTTGTAAACAACCTCTATGT
TCTTAGTGGCAATGAGAAAAGAAACTGCCAGCAGAATGGAGAGTGGTCAAGGAAACAGCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTGACCTGGTGAGAAGGGAGAGTCTCCGATGCAAGGGAGACACCAATTACACCAGCTAT
ACTCAGGGCCTTCAGCAAGCAGAAACTGCAGAGTGCCTACCAAGAAGCCAGCCCTCCCTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTCATCTCACCTTCTACCGCCGCTGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCCTGCATCCCTATGCGGGAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCCATCTACAGGAGGACAGCGGGGTGCATGACGGCAGCCATACAAGGGAGC
GTGTTCTAGTGCAGCGGTGCCCTGGTGAATGAGCGACTGTGGTGGCTGCCACTGTGTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGGAAATTCTACCGGGATGACGGGAGAGAACCATCC
AGAGCCTACAGATTTCTGCTATCATTGCACTCCAACTATGACCCATCCTGCTGATGACATGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCAGTCCAGGCCATCTGCCCTCGCCAGTCGGGATCTCAGCACTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGCTCTGGCAGACGTGAGGAGCCCTGGCTCAAGAACACACTGCGCTCTGGG
TGGTCAGTGTGGTGACTCGCTGCTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTCTGATATCTGCACTGCAGAGACAGGAGGACATGCCCTGTGCTTCCGGGA
CGAGCATCCTGAGGCCAGCGTGCATGTGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATGAAAACCATGTCATGCACTCTTGAGAAG
TGTTCCTGTATATCGCTGTACGTGTCAATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGGCTGTGAACTTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACTCAGTGAAGGGTCAAGTACCTCCATTGCTGGTAGGCTGATGCCGCCCTCCA
CTACTAGGACAGCCAATTGAAAGATGCCAGGGCTTGCAGAGTAAGTTCTCAAGAACAGACCATATACAAACCTCTCCA
CTCCACTGACCTGGTGTCTCCCCAACTTCAGTTACGAATGCCATCAGCTGACAGGGAAAGATCTGGGCTTCTGAG
GCCCTTGTGAGGCTCTCAAGTTCTAGAGAGCAGCTGCCCTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGGCCACAGTACAGTCTGGTCTTTCTCCCTCCTGATACACATTAAATAAGGGTTGGCTCT
GAACTAACAA
AA

FIGURE 38

MELGCWTQLGLTFLOLLLSSSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGORCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSNE
KRTCQONGEWSGKQPICIKACREP KISDLVRRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLOQECISPFPYRRLGSSRTCLRTGKWSGRAPSCI PICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREDETIQSLQISAIILHPNYDPILLDA DIAILKLLDKARISTRVQPI CLAASR
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRSGVVS VVDSLLCEEQHEDHGI PVS VTDNMFCA
SWEPTAPS DICTAETGGIAAVSF PGRASPEPRWHLMGLVWSYDKTC SHRLSTAFTKVL PFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCCTACATCCTCATCTGAGAATCAGAGAGCATAATCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCATAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTG
TGGTCTTGCTTGGCTCAGTCAGTCATAACTACATTGACAATGTGGCAACCTGCACCTCCTGTATTCAAGAACTCTGTA
AAGGTGCTCCCACTACGGCCTGACCAAAGATAAGGAAGAGGCCCTACAAAGATGGCTGTCCAGACGCCGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCCTCCTTAATGACAGACGCCCTGGCTAGACA
ACCCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGGCCAGCAATCAGCCAGTGGACTCTGCCGGAGCAACCGAACTA
GGGCACGCCCTTGAGAGATCCACTATTAGAACGAGATCAATTAAAAAAATTCAGAGCTTGGTGGAGTGTCTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCACCAGGGCAGGGAAATTCTGAAAACACCAACTGCCCTGAAG
TCTTCCAAAGGTTGACACCTGATTCCAGATGGTGAATTACAGCATCAAGATCAATGAGTAGATCCCAGTGAAA
GCCTCTCATTTAGGCTGGTGGAGGTAGCGAAACCCACTGGTCCATATCATTATCCAACACATTATCGTATGGGG
TGATGCCAGAGACGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCACAGGGATGGACATCAGCAATGTCCTC
ACAACATACGCTGTGCGTCTCTGCCGGCAGCCCTGCCAGGTGCTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAAACATGGACAGCCCCGGATGCCCTACAGACCCCGAGATGACAGCTTCTATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGTTGGAATAAAACTGGTGGCAAGGTGGATGAGCCTGGGTTTCTATCTCAATGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGGAGGAGATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAAGTGCGGCTCATCTGATTCCAGGCCAGTGAAAGACGTGTTCACCTCGTGTCCCGCCAGGTTGCCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAGCAATGCCAGCTGGTCCCCAGGGCAGGGAGAGCAACA
CTCCCAAGCCCCCTCCATCTACAAATTACTGTCATGAGAAGGTGGTAAATATCCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTCGCAGGGGAGCATCACATAGAGAATGGGATTGCTATCTATGTCATCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAGAATAAAACAGGTGACATTGTTGAATGTTGAGTGGATGGGTCGAACGTACAGAGGTGAGC
GGAGTGAGGAGCTGGCATTATTGAAAAAGAACATCATCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACACAAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGATTACACCGTGCTGTATAACTGTAAGATATTGTATTACGAAGAACACAGCTGGAGTC
TGGGCTCTGCAATTGAGGAGTTATGAGAATACATGAAACAAACCTTTTCTCAATCATTGTTGAAGGAA
CACAGCATAATGATGGAAGAATTAGATGTGGTATATTCTCTTGCTGTCATGGTAAAGTACATCAGGAATGA
TACATGCTTGCTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTCTGGCTGGCACTT
TTTTATAGAATCAATGATGGGTAGAGGAAACAGAAAAATCACAAATAGGCTAAGAAGTGAACACTATATTATC
TTGTCAGTTTATATTAAAGAAAATACATTGAAAAAGTCAGGAAAGTATGATCATCTAAAGCCAGTT
ACACCTCAGAAAATATGATTCCAAAAAAATTAAACTACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAACAACTAAAATGATTGATTGTATACCCACTGAATT
CAAGCTGATTAAATTAAATTGGTATATGCTGAAAGTCTGCCAGGGTACATTATGGCATTAAATTTACAGCT
AAAATTTTAAATGCATTGCTGAGAACGTTGCTTCATCAAACAAGAATAATTTTACAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPEAAHLIQASERRVHLVVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNIQKDGPESLGMTVAGGASHREWDLPIYVISVEPGGVISRDRG
IKTGDILLNDGVELTEVSRAVALLKRTSSSIVLKALEVKYEPMQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRIITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACACGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCCTGACAGTGGTTGGTGGC
CACCAAGTAACACTTCGTGGGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCTCATTTGGGAGGGAAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACGTGCTTCTGTGTCTCCTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTCGTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
AAAAAGTTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACAGTGGATATTGG
GGGTGTTACTGCCCTAACGAGAGACAGTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGAT
GGGGAGGCAGACGATGACCTCAGACTCAGGGTGAGCTCCAAGAAATGAAAATTCCCGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTTACACCAAGTGTCAAGACTGGAGAACAGATGGTTGAGTAGTTGTTCTTAAAT
TAGTATCTGTGGAACACAATCCTTATATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGAATGGAAAGAACTGATTCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTGTATT
TAGCAGAGCTCCTGGTGAATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTAGTCATTGAT
CATGAGGGTTAAATATTGTAATATGGAACTTGAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCAATTCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGGT
GAGAAGGCCTCACAAAGAGGGAGAAAGGCAGCAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGAGGACCTGCCAGGTATGCCCTCCAGTGATGCCACCAGAGAAATACATTCTTATTAGT
TTTAAAGAGTTTGTAAAATGTTGACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTA LSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTCCCACAACAG
ACGGGACAACTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCAT
GTTCCAGAGGCGAAGGAGGCGAAGACACCCACTTCCCCATCTGCATTTCCTGCGTGCGGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACTGCCCTGCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCCAGAACATAGGTCTTGGAAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAA

FIGURE 44

MALSSQIWAACLLLLL~~LL~~ASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHR~~S~~KCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAGCAAGTTGACTCTATTGCTG
GACCTTCAACACAACCCCTTGTCAACCATACAGCCAGAAGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGATATACAGCTCATCAGCAGGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGTATGACCTCATCTCGTTGCCAGGAACCCCTGTCAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCTGTTGGTGCCCTCCTGCTCAGTCTTTGACTGGCTATTCCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTCGGG
AAACTCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCTAACAGAACAGTCCAGCAAATACGGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACCTCCCTAACGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPIALKLCEGAADDPDSSMVLLCLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTCAGGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATTCATCCAGAACCTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTCTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGAATAAAATGTAAGTATCAGTAGTTGAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMILISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCCTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCAGGGGCTCGGCCATTGCTGGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGACAGTCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTCGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCTGGAGGCCTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCC
TTCTGATCAGGAGGCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHQSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGGACAGAGGGCACAGAGACGCAGAGCAAGGGCGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTCCAG
GGGCCCCCTGGCCTGCCCTGCTGGCCCTGCGCTGGCAGTGGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGACAAATATTGGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAAGAGAAGCAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGC
AGCAGATGCTTGGCAACAGGTGGAAAGCAGCCATGCTCTGGAAACACTGGCAGCAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCCGGCTCTGGCAGGG
GTGCCTGGCCACAGTGGCTTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTAAGGTGG
CCTTGGAGGCCAGGGCAAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGGACCAACACTCAGGGAGCTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGGAGGTGACAGCGGAGT
CCTCCTGGGATCAGCACCGGCTCCTCTCCGGCAACCACGGTGGAGCGGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCGGAGCGGGGAATCTGGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTCTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACCTGAGACGTCTGGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACCAAGAGCTCTC
GCATCCCGACCTCCAGACAAGGAGCCACAGATTGGATGGAGCCCCCACACTCCCTCTAA
AACACCAACCCCTCATCACTAAATCTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCAAAA
AAA
AAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGP
GHSGAWETSGGHGIFGSQGGLGGQQGQGNPGGLGTPWVHYPGNSAGSFGMNPQGAPWGQGGNGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGSGSQSGSSGSGSNDDNNNGSSSGS
SSGSSSSGSSGGSSGGSSGSSGSSGSGSSESSWGSTGSSSGNHGGSGGGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCTGGCTACTCCTGCTGCTGGTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTGGACCTATGCCCTCTATAACAACGCGCCGGCTCCAGTGTTC
CACAGCCCCAAACCGAACGTTGGGTCACCTGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCC
CATCCCCCTCATGCTTATGCCACCCGACACCCTGGCTATCACCAATGCCCTCAGCTGCC
TTGCACCCAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCT
AGTGGCGGTGACAAGTGGAGCCGACCGTCGGATGCTGACGCCCTGGCTTCCATTCAACATCC
GAAGTCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCACTTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAACGCAGCATATCCTCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTCACAGAC
GCTGTCATCCGGAGCGCGTCGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTCATTGATGTCGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCCTGGTCTGTACAACCTTGCAGGCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCAGTCTAAAGAGATTGAATGGGACGACCTGCC
AGCTGCCCTCCTGACCATGTCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCCCTGGGATCCTGAGGTCTACGACC
CCTTCGCTTGACCCAGAGAACAGCAAGGGAGGTCAACCTCTGGCTTTATTCCCTTCTCCGCA
GGGCCAGGAACTGCATGGCAGGCAGTCGCCATGGGGAGATGAAAGTGGCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCAACTGAGCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGTGGAGCCCTGAATGTAGGCTTGCAGACTTCTGAC
CCATCCACCTGTTTTGCAAGATTGTCATGAATAAACGGTGCTGTC
AAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLVVGSLLARILAWTYAFYNNCRLQCFPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRRLTPAFHFNILKSYTIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYILSHDGRRFHRC
RLVHDFTDAVIRERRTLPTQGIDDDFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPPEVYDPFRFPENSKGKSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACTTACCCGTGTTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTTGC
ATAATTCATGGCCAGTTTATGAAGCTTGAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGTCTTACAGACATGTGCCTTTATCTTGCAGCAATGTGTTGC
GTGATTGAAACATTGAGGGTTACTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCC
TCTGGATGTTGTCCCCTGAATTCCCATGAATAACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCTGCCCCG
TACGTGCCGGTCATCGGGAAAGCCCAGACCGAGTTCCAGTACTTGAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCAGTGGAAAGCAGAAAATTGACAAGCT
GGAGATAAGGACCTTGATGGGAGCTAGACTTGAGAATTGTCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAGAATGATGGACGCAATTGACGCCAGGAGATCATGCAGTCCCTGCGGACTTG
GGAGTCAAGATCTGAACAGCAGCAGAAAATTCTCAAGAGCATGGATAAAAACGGCAGCATGACATCGACTGG
AACGAGTGGAGAGACTACACCTCTCACCCCTGGAGAAAATCCTGGAGATCATCCTCTACTGGAAAGCATCCACG
ATCTTGATGTTGGAGAATCTAACGGTCCGGATGAGTTCACAGTGGAGGAGGGCAGAGGGGATGTGGGAGA
CACCTGGTGGCAGGAGGGGGAGGGCGTATCCAGAACCTGCACGCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCATCCCGCAGCAACACATGGGCATCTGGTGTGCTTCACTCAGATGATTGAGAAGGAGGGCCAGG
TCACCTGGCGGGCAATGGCATACGTCCTCAAATTGCCCCGAATCAGGATCAAATTCTAGGCTATGAGCAG
ATCAAGGCCCTTGTGAGTGGACAGGAGACTCTGAGGATTACAGAGGGCTTGTCAGGAGGCTTGGCAGGGTCCCTGGCAGGGCC
ATCAGGCCAGAGCAGCATCACCAATTGGAGGCTGAGACCCGGATGCCGAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGGATCTGGCAGAGAGGGGTTGGCCGCTTCTACAAGGCTATGCCCCAACATGCTG
GGCATCATCCCTATGCCGGCATGCACCTGCACTGAGACGCTCAAGAATGCTGGCTGCAGCACTATGAGCTG
AACACGGCGGCCGGCGTGTGCTGGCACCATGTCAGTACCTGTCAGTACCTGTCAGGCTGGCCAGCTGGCCAGCTAC
CCCCCTGGCCCTAGTCAGGACCCGGATGCCGCAACCTCTATTGAGGGCCTGGAGGTGACCATGAGCAGCCTC
TTCAACATATCCGGACCGAGGGGCTTGGGCTGTAAGGGGCTGGCCCAACTCTCATGAAGGTATCCCA
GCTGTGAGCATCAGCTAAGTGGCTACGAGAACCTGAAGATCACCTGGCGTGCAGTCGGGTGACGGGGAGGGC
CGCCCGCAGTGGACTCGTGTGATCTGGCCGCAGCTGGGTGTCAGGATCTCATTTGTGAATGTGCCAACACT
AAGCTGCTCGAGCCAAGCTGTGAAAACCCTAGACGCCACCCAGGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCCCTGTCACATGTAAGGACAGGACCTCTGTCAGCTGGCCACGGGCTTGGGAGGCTTGGGAGGCTCAGCAGG
CTCCGGCTCACATGTCAGGACAGGACCTCTGTCAGCTGGCCACGGGCTTGGGAGGCTTGGGAGGCTCAGCAGG
TAGTTCTCCATTCAACCTTGCAGCAGCTGGCCACGGGCTTGGGAGGAGGCTACAGCCCACATCCCACCCCTGTC
CCTCTGCTGCCTGCTGTGTCAGGTAAGGTGGGAGGAGGCTTGGGAGGAGGCTTGGGAGGCTTGGGAGGCT
ATAATCATGATGAAAGGTGAGGTACGTGTCAGCTCCAGGCTGACTTCCAAACCTACAGCATTGACGCCAAC
TGTGAAGGAAGGAAAGGATCTGGCTTGTGTCAGTGGCTGAGCTGGCATCTGGGCTTGGGCTCTGGGAGG
CTTGGGAGTGCAGGGGCTGGGCTGGCTGCAAGAGGCAACTTGGGCTGAGCTGGGCTTGGGCTCTGGGAGG
GGCCTGGACCCCTGTCAGGATGGGCCCCACCTCAGAACCAAACACTACTGTCAGTGGGCTTGGGAGG
CCATGTTGAGGGCAAGGGCAGAGCTGGTGTGTCAGGAGGGGAGGAAAGGAGGAGGAGGCTTGGGAGG
ACTGTTGGAAAAGGGTTTGTCCAGAAGGACAAGCCGACAATGAGGGACTTGTGTCAGGAGGGAGG
GAGCAGGAGCTGGCTGACTGTCAGAGTGTCTGACGCCCTGGGCTTCTGCAACACCCAGCAGGGGCCAGC
GGGACCCACATTCTGTCAGTGGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
AACCTTGAAGGTGGAATCCAGTTTCTGCTGAGGTTCTTATTTCACTCTTCTGAATGTCAGGCAG
TGAGGTGCTCTCACTGTAAGTTGTTGGGCTGGGAGGGCTGGAGGAGGGTGGGGGCTGGCTCCGTCCTCC
CTTCTGCTGCCCTGCTTAACAATGCCGCCACTGGCAGCTCACGGTTGCACTTCAATTCCACCA
TGAGGAATCTCAATAGGATGCAAGAATGCAAAATTGTTATATGAAACATATACTGGAGTGTCAA
CAAATTAGAAGAATTGGACGTTAGAAGTGTCAATTAAAGCAGCCTCTAATAAAGTTGTTCAAAGCTGAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEESTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGTTAGGGAGACACTCCATCACAGTCACACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTGAACCTGACATCAAACTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGGCCAGCAGCAGTGTGCTGATCAAAGTGTAGTT
GGCAATGCCCTTTGGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGAAGGGAAATGCTAACCTTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGCTGGGATCCAAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGT
CTCTTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTAGCCCTTACCTGATGCTAAATTAAT
GTGCCTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCAC
CACCAGATATGACCTAGTTATATTCTGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAAACAAGAGCAAGAAACAAAAGAAGCCAAAGCAGAAGGCTCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTCACTGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCGT
CACCTGGGAGTGAGAGGACAGGGATAGTGCATGTTCTGTCTGAATTTTAGTTATATGTGC
TGTAAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCCTAAGACGCTGCTAAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAATAAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSEVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNASKL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAACATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTGGTAGT
GGGCTACCCCCTGGTGGTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCTCCAAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCACGGCTCTCAGCCCTGTGGAAATTATCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGCCC
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTATGACCAACCATCTGTGGAGAGTGGAGTGCACCAAGGGACCTT
TCCTGGCTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTCTGTCCCTAGAAGAAC
ATTCTCCCTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATAGCCACCA
AAATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCAGCAGCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTAGTGTCTCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCGCCATCCTGTCACAGGCCATTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGCTATGGACCAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCATTTTTTC
ATTGGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSPGTGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLVLTAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTGCCCCGGAGCATCGGTGTTGAGGAGAAAGTTCCAAAACTTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTCCTCCACTGGCCCTCTAACTCTGAACATCCGCAGCCC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGCTGAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACCTCTTACCTCTCCAGTGCGCGGCCCGCTCCGGCAGTGGCCTT
TGCCTGGGGAGTCTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGAGCCGGGGAAAATCCTTCCAAACG
CCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATACCCCTGGGTACCCTGAATCCCAGTG
TGTCCTGGGGAGGTGGAGGCCCTGGGACTGGTGGGGAACGAGGCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAGG
GCAGCTGGGGAAATATTCATCTACCCAGGTATCAATAACCCATTCCTCTGGAGTTCTCCGC
CCTCCTGGCTCTTGGAAACATCCCAGCTGGCTCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCCAACATTGGAGTTAGAGTCCTGCTCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCACTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCATTATGAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLLVLCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRLPRSNSILGAGGKILSQRP
PWSLIHRVLPDHPWGLTNPSVSWGGGGPGTGWGTRPMPHPEGIWINNQPPGTSGWNINRYPGGS
WGNINRYPGGSGWNINRYPGGSGNIHLYPGINNPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGGCCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGCTCCTCCTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTTCTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCCGACCCGTCTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAAATTGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATGCGGTAGGGGGCGAGCGCGAGAACCCCC
TTCCTCGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCGGGCTGGGCTGCTTCTGGCG
CTGGGCGCTGCCGTTCTGCTGGCCCGCTGGGGCGAGCCTGGGGCAAATACAGACCACCTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTGGCTGCCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGGGAAGCTTCGGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATTAGGTCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGGTA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGGTA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATATTTATATAAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIOTTSANENSTVLPSSSSDGTLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKQTEGTYRPSSEEQFSAEAARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFKKNVSIILIPENWKENPQYKRPKHENKHADVIVAPPTLPGRDEPYTKQFTECGEKG
EYIHFTPDLGGKKQNEYGPPGKLFWHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVT
PPPPPFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLLOQTVEENGWSWGMVHFSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVICELHSQLDGSEVLLLTGDGEDNTASSCIDEVKQSGAIVHFI
ALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDA
FGALTSGNTDLSQKSLQLESKGTL
LNSNAWMNDVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKAN
PETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYV
FVLGANVTA
FIESQNGHTEVLELLDN
GA
GADSFKNDGVYSRYFTAYTENGRYSLKVRA
HGGANTARLKLRP
PLNRAAYIPGWVVNGEIEANPP
RPE
IDE
DTQTTLED
FSRTASGGAFVV
SQV
ESLPLPDQYPP
SQITD
LDATV
HEDKI
ILTWT
APGDN
FDVGKVQRYIIRISAS
ILDLRDSFDDALQVNT
TDLSPEANS
KESFAFK
PENISE
ENATHIFIAI
KS
IDKSNL
TSKV
SNIA
QV
TLF
IPQAN
PDD
IDPT
PTPT
PTP
DKSH
NSGVN
ISTLV
LSV
IGSVVI
VN
FIL
STTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCCTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTGCCCGGGCAGGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCCTCTATTCAAGGAAAGACGCCAAGGTAAAGGGCA
GAGGAGCAATGATGTAGCCACCTCTAACCTCCCTCTTGAAACCCCCAGTTATGCCAGGATTACTAGAGTGTCA
ACTCAACCAAGCGCTCTCGGCTTAACCTGTGGTTGGAGAGAACCTTGTGGGGCTGCCTCTTAGCA
GTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAGAAAAGGAAAGGTCCCCTCTGCTGTTGGCTGCCACATCAGGAA
GGCTGTGATGGGAATGAGGTAAAAGGAACTTGAGATTTCACTTCAGTCATTGCTCTGCCAGAATCATCCTTAAA
AGTAGAGAAGCTGCTCTGTGTGGCTTAGCCAGGGCAAGACTCTTCAAGGAGAAATGGATGCAAGCAGCTC
CGGGGCCCCAAACGATGCTCTGTGGCTAGGCCAGGGCTCCGTGGGGCCGGCTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTATTCTGAATGATGATGGTCGCCGGGGCTGCTTGCGTGGATTCCGGGTGGT
GTTTGCTGGTCTCTGCTGTGCTATCTGTGCTGTACATGTTGGCTGCAACCCAAAAGGGTACAGGCTCAGGAG
CTGGCACTGCCAGGGCAACAGCCACGGGAACAGGGTACAGGCCAGCTGCAAGGAGGGTGA
CGCAACTACGTGAGCACCTGAAGCGGAGATGCCAGCTCAAGGAGGGTACAGGCCAG
AATGGCAGTACCAAGGCCAGCGATGCTGCCCTGGCTTGACAGGAGCCCCAGAGAAAACCCAGGCCACCTC
CTGGCTTCTGCACTCCAGGTGCAAGCAGGGTGAATGCTGGCTCAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTCGATAGCTTACTCTACAGAAGGTGTACAGCTGGAGACTGGCTTACCCCAACCCGGAGGAGAACGCTGTG
AGGAAGGACAAGGGGATGAGTTGGTGAAGCATTGAATCAGCCTTGGAGACCTGAACAAATCTGCAGAGAACAGC
CCAACTACCGCTCTAACGGCTGTGATTTCAAGGAGTCAAGGAGAACAGAAGGGACAAAGGGACATTGTAT
GAGCTCACCTCAAAAGGGACCAACAAACAGGATTCAACAGGCTCATCTTATTGACCAATTGACCCCATATGAAA
GTGAAAATGAAAAGCTAACATGCCAACACGCTTATCAATGTTATGTCCTAGCAAAAGGGTGGACAAGTTC
CGGCAGTCATGCGAGAATTCAAGGGAGATGTGCAATTGAGCAGGATGGAGACTTCAGTGTGTTACTTTGGG
AAAGAAGAATAATGAAGTCAAAGGAATCTTGAAACACTTCCAAGCTGCCAACACTTCAAGGAACTTACCTTCATC
CAGCTGAATGGAAATTTCAGGGAAAGGGACTTGTGATTTGGAGCCCTCTGGAGGGAAAGCAAGCTCTTC
TTTTCTGTGATGGACATCTACTTCACATCTGAATCTCAATACGTCAGGCTGAATACACAGCAGGGAAAG
GTATTTATCCACTTCTTCAGTCAGTACAATCTGCCATAATATACGCCACCATGATGCACTCCCTCCCTGGAA
CAGCAGTCGGTCAAAAGAAGGAACCTGGATTTGGAGAGACTTGGATTTGGAGCTGCAAGTATCGTCAGAC
TTCATCAATATAGTGGTTGATCTGGACATCAAAGGCTGGGGAGAGGATGTGCACTTATCGCAAGTATCTC
CACAGCACCTCATAGTGTGACGGACCTGTGCAAGGACTCTTCACCTCTGGCATGAGAACGGCTGATGGAG
CTGACCCCCGAGCAGTACAAGATGTGCACTGCCAGGGCATGCCACAGGAGCTGGCATGTC
GTGTTAGGACAGAGATAGGGCTACCTTCGCAAAAGAACAGAACAGAACAGTAGACAAAAAACATGA
AACAGGATTGTGGGGACACTTTTCTTCCTTGCAATTACTGAAAGTGGCTGCAAGCAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAATTGGACTGATGGCTAGAGATGAGAAAGCCTGGGCTGATTCTCTGTGTTGGCTTTTACAACAGA
AATCAAATCTCGCTTCTGCAAAAGTAACCCAGTGTGACCCCTGTGAAAGTGTGACAAAGGCAATGCTTG
AGATTATAAGCCTAATGGTGGAGGTTGATGGTTACAATACACTGAGACCTCTGTTGTTGTGCTCATG
AATATTCAATGTTAAAGAGCAGTTTGTAAAATTCATTAGCATGAAAGGCAAGCATATTCTCTCATATGAATGA
GCCTATCAGCAGGGCTAGTTCTAGGAATGCTAAATATCAGAAGGCAAGGAGAGGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAAATAAAATGGACAGGAAAGAAGAACATAATATGTCATATTCCCCAAGAT
TAACCAAAAATGCTTATCTTTGGTCTTTAACTCTCTGTTTCTTATTTAATTTAAAGTGC
TTTTCTCTGTGAGTTAGTCTGTTTACCTGTTTAACTCTGCTTACAGCTGGGATGATGTCACCTAAAGTCA
ATTTTATATTGTTAAGAGATACCTTGAGATGCAATTGAGACTTCAAGCTGCAAGCATCAAATTGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGCAAGGACTGCAATGTCAGGCAATTGAGACATAGGGAAGGAATGGTTGACT
AATACAGACGTACAGATACTTCTGAAGAGTATTCTGAAGAGGAGCAACTGAAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGTTACAGAAAAGGAAACTCATTGAGACTGCTGATGTCACCTAAAGTCA
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAAACCAAGTACCTGATGCA
AAAACAGGGTGCTCTCTGCTTCTGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATTGT
GAAAGATCAATCCATGCGAGAATCTAGTGGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGTGGAG
TAACTGAATTATTTTAAATTAGCAGTTACTCAATCAGGAGATGCTGAAATGCA
CAAACATTTTAAATAATACAGTAAACATAAGTGGTTCTCATTGATGCA
ATGCTGAGCAGTAAATTATCTCTTGAGTCCTGCTGTTGCTCAAGTAAACCTATGTTAAAAGCTTCAAGAAC
ATTCAAGCTGGTGTGTTAAAATGCAATTGATTTGATGTTGAGTTATGAAATTAAACACAGG
CCATGAATGGAAGGTGATTGACAGCTAATAAATGATTGTTGAGTGA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINIVPLAKRVDKFQFMQNREMCEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQSQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTQCQYRSDFINIGGFDLIDIKGWGGEDVHLYR
KYLHSNLIVVTPVRGLFHLWHEKRCMDELTPEQYKMCMSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCTGGGATCCA
GAAACCCATGATACCTACTGAACACCGAATCCCCTGGAAGGCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCTAGTCCTAGTCCTAAATTCCAGTCCCTGCACCCCTTC
CTGGGACACTATGTGTTCTCCGCCCTCTGCTGGAGGTGATTGGATCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGCCAGCCTCTTACCCGTAGTGT
GGAAACAAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCGTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCTACCCGTATCTGGGTGACTTCCCGAAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCAGGGGGTCAGAACACAGATCAACAG
TGAAGCCACATTGCAAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGGCTCAGGGCCTGGCTGTCCTGGCATCCTAATTGAGGTGGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAAGAGCTGCTGGGACAGTTTTATAGAAGGTCCAG
ATTCAATGAAACAGCTGAAAAGCTCAGGGACATTGTTCTCCACAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTTGCTTCTTCA
TCCAAGCAGGATCCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTCTGCCTCTCTGGCTGTTATTCATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGCTTCACCTCAGCACAAGCCACGACTGAGGCATAAAATTCCCTCAGATA
CATGGATGTGGATGACTTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGGTAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAAGACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTTGCAAGGGAAAGTTGG
ATATACCCCAAAGTCCCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGATATTTGATCAATATAATTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACACTGCAAACATCACCTCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGCTTCAGGAGTTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGGCCCACCAGAGGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTGTCCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGAATACTAAACAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTGCCTATCTATTACCGTGTGTTCTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTATGGAAATGAATTGACAAAAGATTCTTGTGCCGTGCTGAAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGTTGCGATTGATGAAATTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGACTCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGCCCGAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTCAGTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCCTCGCTGTCAGCTCGACCAGGATTAGAGGGCTGCG
AGCCTCTGAGGGGATGGGCTGGAGAGGGTCTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCTATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTCCTTGCCTTGTGAAAGTTTCCAGTTGTCAGTGT
TCACCCCTTGATCCCAGCCATAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCAGTGT
CTGTGAGAATTACTTATTCTTCTATTCTCATAGCACGTGTGATTGGTTATGCATGTA
GGTCTTAACAATGATGGTGGCCTCTGGAGTCAGGGCTGGCCGGTTGTTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGAGAATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSILQEEVSTQGTLLESQA
ALAVLGQPQLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSILSSFDQDS
EGCEPSEGDSLGEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMAN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCAC TTGCTGCCCTGTGACAC
CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCCTCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGGCCACCATCCGCATGGACACCAGTGAAGTGGC
CCCACCCGCCCTGGTCCCTCAGTGAUTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAC TGCTGTA
TAAGCTCTCCTCCCTGGTGAACGCCCTAGCTAACAGGTCAAGCAGGTCAAGCAGGTCTGGAGTTGACCTCTGTA
CCAATCTAGTAAAAAACCAGCTGTGTCCCGTGATCGAGGGCTTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCAAGCTCTACCTGGGGCCAAGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGAGCTGCACTGGCTGCTGTGCTCTCCAGAAGA
ATTCACTGGCTCTGGACTCTGTGCTTCTGAGAGTGCCATCGCTGAAGTCAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTAAGATCTTAACCTCAGGAC
ACTCCCAGTTTTATAGACCAAGGCCATGCCAAGGTGCCAAGTGGCCAACTGATCGTGCTGGAAAGTGTT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTTCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCA
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCTGTCTCCAGTGAAGAGACTTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLQLQVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIASSEAQFYT
KGDQLILNLNNISSDRQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTGTGGGCTACATCCTAGGCCTCTGGGGCTTTGGGCACACTGGTGCCAT
GCTGCTCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCCAGTCTTCTGCCAGGAATCCC
GAGCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCCTGGGATTC
ATTCCCTGTTGCCTGGAATCTCATGGGATCCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTCTTACTTGGGATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCTGCTCATCCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCAACCTTGCACAAAGGAGCTCTCAAGGCCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGGCCAGACTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCGAGGGCCACAGGTGAGGGACACTACCAGTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGAACAGCATGCAGGTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTGCTGCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGGCCCTGGTTACCTGGACTCCATCCCCAACCCACTAATCACA
TCCCACGTGACTGACCCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGGCCCTGGAACCTCCACTCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSIASSLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGI
LGFIPVAWNLHGILRDFYSPLVPSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCCACCGCGTCCGCCTCTCCCTCTGCTGGACCTTCCCTCGTCCTCCATCTCTCCCTCTTC
CCCGCGTTCTCTTCCACCTTCTCTTCCACCTAGACCTCCCTGCCCTCCCT
GCCCACCGCTGCTTCCCTGGCCCTCTCGACCCCGCTAGCAGCAGACCTCTGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCGTCTTCTGCTCCCTCCGACTCCGCTCCGG
ACCAGCGGCCTGACCTGGGGAAAGGATGGTCCCGAGGTGAGGGTCTCTCCCTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGGCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTCAGCCTGTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCAACCGAGTGTGTCCCTGAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCCGAACCGAGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAACGCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTCATCCCTGCCACT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAACATAAGAAAGCCT
GTGTGCATGGCGGAAGACGTACTCCACGGGAGGTGTGGCACCCGGCTTCCGTGCCTCGC
CCCTGCCCTGCATCTATGCCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAG
ACAAAGCAGCCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCCGTGCTTGCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGCCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCATGCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATTAATAA
TAAGAAGTTGCATTACCCCTAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFS AHELF
PSRLPNQCVLCSCTEGQIYCGLTTCEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGT PAPTGLSAPLSFIPRHFRPKAGSTTVKIVLKEKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPPRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCCGCTACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCCTTCTCTAATCCATCCGTACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGCCATTACAGAACACATCCTGGCTCTCATGCTCAGTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGCCTGTCCAGGCCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCTGTCCTAAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTAGCGTGGTCCACCTCTACAGGGCAGGGAGGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGGGAGGGGCGATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTCCCAGTCTACTACCAGAAGGCCATCTGGGAGCT
ACAGGGTGCAGCACTGGGCTCAGTTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTTCCACAAGGACAGGATTGTC
ACAGACTCCAGGACAAACAGAGACATGCATGGCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGGGCATGCTCATCTGAGCCGAGGGTGGAAATCCAGGGTACAGATAG
GAGATACTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGAATACTCTGCTGTGGCTA
TTTTTGGCATTGTTGACTGAAGATTTCTTCCAAATTCCAGTGGAAATCCAGGCGGAACGGACT
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCCGAAGCTCTGCGTTCTGATCTGAAAATGTAACCCATAGAAAGCTCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCGTTTATCAGCGTCTCCCCAGGACCCACCTACAAAATAGGGTCTCTGGACTA
TGAGTGTGGGACCATCTCCTCTCAACATAATGACAGTCCCTTATTATACCCCTGACATGTCGGTTG
AAGGCTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTCTGGCAAAGGCCCTCTGCAATCCCAGAGACAAGCAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCCTCCCCAGGGTGAATGTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAGGTGGCTTCCA
GATGAAGGGGACTGCCGTGTCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATGGCTAAGTGATCTGAAATACCACCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCAATTATATTACACTTCAAGTAAAAAA

FIGURE 84

MALMLSVLSSLKLGSQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRRAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEQELDWRKKHGQAEQELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTACCCCTCGCATGGCTGGATT
TACCTGGCCCAGTAGTCATGGCTACTGGTCCGGAAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGCAGTGTGGAGGAGACTGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGATGCCGGAGA
TACTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCAACCCGCTCCTCGGTGCTACCCCTATCCCACA
GCCCGAGGACCATGGCACCAGCCTCACCTGTCAAGGTGACCTCCGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTAACGTGCTCACCGCCTCAGAACCTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCAGAGGGCCAGTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCCCTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCACATCAGGAGTGAETCAGGGGTGGTGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCCCTGCCTCATCTCGTTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGAGATACGGCATAGAGGATGCAAACGCTGTCAGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTGGCAGAACAGACTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGAGAGGAGGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAGAACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAAGACAGACAAATTCTTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHVPSCFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTSTVLGNSSL
SLPEGQSLRLVCADVDAVDNSNPPARLSLSWRGLTLCPSQPSNPVGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEHELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGGC
CAACTACAACACCTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCCSTAT
GGCCAGCGGGATTCACTGCGGGATTGTTAGTCAGGTATTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCGTGAAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TGAGAGTTTGTGGAGGGACCCAGACCTCTCCTCCAAACCATGAGATCCAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQGSKADYPFGDGNWANYNTFGSAAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTNDNGPVIPVYDFGDAQKTASYSPYGOREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGG~~GACTTCAGGAGTCGCTGTCTGAAC~~TCCAGCCTCAGAGAC~~
CGCCGCC~~CTTGT~~CCCCGAGGGCCATGGGCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCC~~TGACG~~
CTCCTGGCGCATCTGGTGGTCGT~~CATCAC~~CTTATTCTGGTCCC~~GGGACAGCAAC~~ATACAGGCTG
CCTGCC~~CT~~TCACGTCACCCCCGAGGGAGTATGACAAGCAGGACATT~~CAGCTGGTGGCCGCGCTCT~~
CTGTCACCCTGGGC~~CT~~TTGCAGTGGAGCTGGCCGGTTCC~~CTCAGGAGTCTCCATGTTAAC~~
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCC~~CTGT~~CTTCTTCAT
ATT~~CGAGCGTTGGGAGTGC~~ACTACGTATTGGTACATTTG~~TCTGCAGTGCCCTCCAGCTG~~
TCAC~~TGAAATGGCTTATT~~CGTACCGT~~CTTGGGCTGAAAAAGAAACCC~~TTCTGATTACTTCA
TGACGGGAAC~~CTAAGGACGAAGC~~CTACAGGGCAAGGGCGCTCGTATT~~CCTGGAAGAAGGAAG~~
GCATAGGCTTCGGTTTCC~~CCCTCGGAAACTG~~CTGCTGGAGGATATGTGTTG~~AATAATTACG~~
TCTTGAGTCTGGATTATCGCATTGTATTAGTGC~~TTGTAATAAAATGTTGTAGTAACA~~
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGGAGGAAGGATGCAACTGAC
TCGCTGCTGTTCTGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCGGCTCAGAGGACCCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAACGTCAGTCCCAGGGCATGGCAATTCCACTCTCCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTCATCGAAGCCAAGGCCTCCAAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGTATAACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGGC
TGGAAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAAGGGTTGGGCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGAGCCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGG
TCCCCGAGGCCTGTGGCAGGCCATCAGGCTTGGCCAGGGCAGGGCAAGTCTCAAGTGGCTAAGC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGAACAGGGAGGGAGATTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGTGG
GCCAGAGGAGCTCCAGCCCTGCCATGGCCAGGGCTTGAGCCCTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGCTTGACAGATTGACCACATGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCTTAAGCTAACAGACAGGAGGATTGTGGCTCCACACTAACGGCCACAGCCATC
CGCGTGTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCTGGAGCATCCATGTCCCG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACCGGGTTCTCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCCGCAGAGCATGTGCTGGATCTGTC
TGTGTGTCTGTGTGGTGGGGAGGGAGGGAAAGTCTTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGAGCAGGAAATAAGCTTGGCCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNNTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCIHD
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGT~~TTCGGCTGCGC~~CTTCATTGCCCTCGGGCTGCGCTCGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATT~~CGTCC~~CTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGC~~GTTG~~TCTGTCTATATCCAAGAA
ATGTTCCGATTG~~CAT~~ATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAACCC
AGGTGAGACAGCACCC~~T~~TATGCGACTGCTGGC~~T~~ATGTTCTGGCTGGGCTT~~GGA~~ATCATGA
GTGGAGTATTTCC~~TT~~GTGAATACCCTATCTGACTCCTGGGGCCAGGCACAGTGGG~~CATT~~CAT
GGAGATTCTCCTCAATTCTCCTTATT~~CAGCTT~~CATGACGCTGGCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGT~~CAGCCC~~AGACCTTCATAAGTTCTTATTATGGAATAAC~~CTGGCG~~
TCAGCATTATAATCCTGGTGCTCATGGCACCTGGCATTCTAGCTGC~~GGGAGGCAGCTGCCG~~
AAGCCTGAAACTCTGCCTGCTG~~CCAAGACAAGAA~~ACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCCAGCACTTCCAAACCGCAGACTACATCTTAGAGGAAGCACA~~ACTGTGCCT~~
TTTCTGAAAATCC~~TTT~~CTGGT~~GAATTGAGAAAGAA~~ATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFMSGV
FSFVNTLSDSLGP GTVG IHDSPQFFLYSAFMTLVII LLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAAGGSCRSLKLCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGGAGGAGTACGATTCTACTGTTTGCTCTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGTTGGACTCCCTCCCACAAAATGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCTCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAGGAGCACACATGCCATCGAGGAAGGCCACCACAGAAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGATAACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAGAAATTTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTTCAGAAAATATGTTATGTAGAGAAATCCAACCTTTAAAAAA
CAATAATTCAATGGATAAAATCTGTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAATGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTPAGIQRSTHAIIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCCCACCCAGCCCAGCCTGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGCCAGTCTGGCGCTGGAGGGCTGCTCTGACCATG
GTCCCTGCCCTGGCTGTTGTCTCCGCCCCAGGCTCTCCCCAAGGCCAGGCCAGCTGC
AGAGCTGTCTGGAAAGTCCAGAAAATCTATGGTGGAAATTTCCTTATACCTGACCAAGTGC
CGCTGCCCTGTGAGGGGCTGAAGGCCAGATCGTCTGCAAGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTCTGCTGGAGATGCCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCGCTACAGGTACCCCTGGAGATGCCAGGATGGACATGCTTGTGGGCTCAC
AGCCTGTCTTGACCGTGAAGGATGAGAATGACCAGGTGCCCTTCTCAAGGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCGGCCCTGGCATCCCCCTCCTCTGAGGCTTCAGACCG
GGATGAGGCCAGGCACAGCAACTCGGATCTGATTCCACATCCTGAGCCAGGCTCCAGGCCAGC
CTTCCCCAGACATGTCCAGCTGGAGCTGGCTGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCAAGCCCTGGAGAGGACTACCGAGTGTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAGTCTCCATCATAGAGACACCTGGGTGT
CCCTAGAGCCTATCCACCTGGAGAGAATCTCAAAGTCTTATACCCGACCATGGCCAGGTA
CACTGGAGTGGGGTGTGACTATCACCTGGAGAGGCACTCCCCGGGACCCCTTGAAGTGA
TGCAGAGGAAACCTACGTGACCAAGAGAGCTGGACAGAGAACGCCAGGCTGAGTACCTGCTCC
AGGTGCGGCTCAGAATTCCATGGCAGGACTATGCGCCCTCTGGAGCTGCACGTGCTGGT
ATGGATGAGAATGACAACGTCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTCACTGAGACTGTCAGCAGAGGATGCCAGATGCCCGGCTCCCCA
ATTCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGGAGGCCCTC
CAGGTGGACCCCACTCAGGCAGTGTGACGCTGGGGTGTCCCACTCCGAGCAGGCCAGAACAT
CTCGCTCTGGTGTGGCCATGGACCTGGCAGGGCGCAGAGGGTGGCTTCAGCAGCACGTGTAAG
TCGAAGTGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGATTGGGCT
ATAAGCTCCCTGAGGATGTGGAGGCCGGACTCTGGTGGCCATGCTAACGCCATTGATGTC
CTCGAGGCCGCTTCCGCTCATGGATTGGCAATTGAGAGGGAGACACAGAACAGGACTTTG
GCCCTGGATTGGAGGCCAGACTCTGGGATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCAAGTCATGAGGTGGTGTGAGCTGCTAGTGGAGAGACTGATGCCACCCCCAAGTTGG
AGGCCCTGGAGGCCACGCCAGGGTACTGCTAGTGGAGAGACTGATGCCACCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGGCCAGGGCTTCTCTGCTGACCATC
CAGGCCCTGGACCCCACTCAGGCCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCATTGAGAATTCTGGGAGGGTGCCTGGGAGGCCAGGATACGCCCTGACTCTGCCCTGTGCCCTCC
ACACCTACACGGTGTGTTGGAGGCCAGGATACGCCCTGACTCTGCCCTGTGCCCTCC
TACCTCTGACACCCCCGCAAGGACCATGGCTGATGCTGAGTGGACCCAGCAAGGACCCGATCT
GCCAGTGGGACCGGCTTACAGCTCACCTTGGTCCAACCCCACGGTGCACGGGATTGG
GCCTCCAGACTCTCAATGGTCCCATGCCCTACCTTGGCCCTGCAATTGGGGAGGCCAGT
GAACACATAATCCCCGGTGGTGGTCAAGCCACAATGCCAGATGTGGCAGCTCTGGTTCAGTGT
CGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGCCGATGAAGGGCATGCCA
CGAAGCTGCGCAGTGGCATCTTGTAGGCACCCCTGGTAGCAATTAGAATCTCCTCATCCTC
ATTTTCAACCAACTGGACCATGTCAGGAAGAAGGACCCGATCAACCGAGACAGCGTCCCC
GAAGGCAGACTGTCTGATGGCCAGGCACTGCTAGTGGAGCTGGCCCTGGCTCCATCTGAG
TCCCCCTGGAGAGGCCAGCACCCAGATCCAGCAGGGACAGGACAGAGTAGAAGGCCCTCCA
TCTGCCCTGGGGTGGAGGCCAGCATACCATACCCAGGATGTCTGAGGCCCTGGACACCAACTT
TATGGACTGCCCATGGGAGTGCTCAAATGTCAGGGTGTGTTGCCATAATAAGGCCAGAGAA
CTGGGCTGGGCCCTATGGGAAAAA

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVICPPRDPPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYVQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNILLVLAMDLAGAEGGFSSCTCEVEVAVTIDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGDWEPDSGHVRRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPAGATATVVLVERVMPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCEIKFSGEVHTAQSLQGAQPGDTYTIVLVEAQDTALTTLAPVPS
QYLCTPRQDHGLIVSGPSKDPDIASGHGPYSFTLGPNPPTVQRDWRLQTLNGSHAYLTIALHWVEP
REHIIIPVVVSHNAQMWMQQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTLVAIGIFLI
LIFTHTWMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCAACGCC
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAAGCACCAGCTT
TCTCTTCACCTTCAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCCTATGTTGGTCTACTATTGCATTAGAAGCTCAACAAATTCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCCAGCACAGGCCAACACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACATCTCAGGGTCCAGCGTACCTCCAAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGGCCAACACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACAAACTCTGAGTCAGCACACCTCCAGTGG
GGCCAGCACAGCCACAAACTCTGAGTCCAGCACACCCCTCAGTGGGCAGCACAGTCACCAACT
CTGGGTCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCAGCACAGTGTCC
AGTAGGGCCAGCAGTCCACCAACTCTGAGTCAGCACACTCTCAGTGGGCAAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACAACTCTGAGTCAGCACAGTCAGTGTCCAGTAGGGCAGCA
GCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCAAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCCACGACCTCCAGTGG
GCACAGCCACCAACTCTGACTCCAGCACAGTCAGTGTCCAGTGGGCCAGCAGCACACAGCCAC
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTCAGTGTCCAGTGGGCCAGCAGCACACCCCTCAGTGG
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCAGTGG
ACACAGCCACCAACTCTGGGTCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACCTCCATAGTCATCTACTGAGTGAGTGAGGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAATCTCCTCATACCCCTGGCTCGGTTGGCGCGTGGGGCTCTTGTGG
TCTTCTCTGTGTGAGAAACAGCCTGTCCTGAGAAACACCTTAAACACAGCTGTCTACCACCC
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCACAGGCCAG
GTGGAGTCCTAACTGGTTCTGGAGGAAGCAGTACATCGATAGCCATGGAGATGAGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGGAAAGCAAGTGCCTGATTCTTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCATTCTCACCTTCTGCCTTACAGACACTGGAAAGAGAAACTAT
AATCTTGAAGAAGGTATTCTCACCTTCTGCCTTACAGACACTGGAAAGAGAAACTAT
TGCTCATTTAGCTAAGAAATAATACATCTCATCTAACACACAGCAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACCTCAGTTAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKM**Q**KGNVLLMFGLLLH**E**AAATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSG**I**ATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVNTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSEASTATNSESSTVSSGISTVNTNSES
STTSSGANTATNGSSVTSAGSGTAALTGMHTTS**H**SASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP~~G~~GNHGAPHRPRWSPNWFWR~~R~~PVSSI
AMEMSGRN~~S~~GP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCCGTGTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCACGCACCCCACTCGCGTCCGCACGGCTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAGCTGGCAACAAAT
GGATGATGTGATATAGCATTCCAGGGGAAGGGAAATTGTGGTGCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGAATCAT
GGTGTCACTGGAAAGGGATTACTTACTGACTCTGTTGGAGCTTTGGAGCTTGGAGCATTT
TCATGCTGAGTCCCCTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGATTATTGGAGACATGTTGGTAAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCACGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCA TAGGAAATGGAAGGATGACAAGAGCCATTGAGACATGATTGATTACTTTGTG
ATATTCACTCGAACCACTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCAAGGAGACTTT
CCCAGGGAAATCCACCTTCACGTCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGGCCACAAACGGTGGGAGAGAAAAGAGAGAGGCTGCGTCCCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCACTCCACCTGCAAGTCTGAACCTAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTCAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTAAAGGGTATTATAATCACCATTGAACTTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAACCTAGAGCATATTG
GAAATGTTCTAAACCTTCTAACGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTTGGAAAAATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGTGATGGGGTGAATAACTT
GGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIFYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRREGKNLDAVHDITVA
YPHNIPOSEKHLLQGDFPREIHFHVRYPIDLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVFVLQERIFGG
LEIIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCACGGCTCTGCAGCTGGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCACTGTTTCACTAGTGTAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTGGTTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCCTGGTGGGGGAGGACGCCGTGTCCTGCTCCCTCT
TTCCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTCAAGGCTAAGGCTAAAAAACATCACTCCCTCGG
ACATCGGCCTGATGGGTGCTGGTTCACTGAGTTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTCCTCTCATTCACATGTTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCCTCAGGCTGGTCCCCCAGGCCACAGCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCCCTGGCTTCTA
TTTACTCGGGTTACTCTGTGGTGCCTGTGGTGTGTCATGGGATGATAATTGTTCTC
AAATCAAAGGAAATCCAGGCGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTCG
TTTCTGATCTGAAAATCTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTGGAGTGTGTCGGGATGACGTAGACAGGGGAAGAAC
ATGTGACTTGTCTCCAAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACAGGAGTAGGGGTCTTCC
GGACTATGAGGGTGGGACATCTCTTCAATACAAATGACCAGTCCCTTATTATACCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCATATCCAGCATGCGATGTATGACGAGGAAAG
GGGACTCCCATATTCAATGTCAGTGTCTGGGAT**TGAGACAGAGAACCCCTGCTAAAGGGC**
CCCACACCACAGACCCAGACACAGCAAGGGAGAGTGCCTCCGACAGGTGGCCAGCTCC
CCGGAGCCTGCGCACAGAGTCAGCCCCACTCTCTTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCAGCGCAGTCAGCTCCAGATGAGGGGGATTGGCCTGACCCGTGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGAGGAATAGACTCACATTAGGTTAGTTGTGAAAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAAACGAATGTGAATCATGCTTGAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTATATTACATTCCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTACCAAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAAGAGGGAGTAGGATTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACAGGCAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATAATT
AAGATGATATAAACTACTCAGTGTGGTTGTCCCACAAATGCAAGAGTTGTTAATATTTAAAT
ATCAACCAGTGTAAATCAGCACATTAATAAAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKOMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCCLSSGWFPQPTAKWKGPOQDLSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPSPWRLASILLGLLCALCGVVMGMIIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVRLLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGGTGGAACCTGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGCTCTGGCAGTGTGCATTGGA
CTCACTGTCATTATGTGAGATATAATCAAAGAACGACTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAATCTCCATTAAGGGAAAGAATTGTCAGTCTCAGGTTATCAAGTT
AGTCAACAGAACGATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATGTTCAACTTGTACATGAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG
ATCCTCACTCAGTAAAATTAAAAATCAACAAGAACAGAACAGACAGCTATCTAAACCATTGCTGCCGA
ACACGAAGAAGTAAAACCTCTAGGTAGAGTCTCAGGATCGTGGTGGACAGAAGTAGAAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAGCCACATGGC
TTGTGAGTGTGCTCACTGTTACAACATATAAGAACCTGCCAGATGGACTGCTTCCTTGGAGTAACA
ATAAAACCTCGAAAATGAAACGGGTCTCCGGAGAAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGAGAGCTTCTAGCCCTGTTCCCTACACAAATGCACTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTAGTGTGATGTTGTGACAGGATTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACATTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACCACTGGTAGTTCAAGTGTAGAGATATGGTACCTTGCTGGAATAGTG
AGCTGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTTACTAGAGTTACGGCCTTGCAGGGACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTGTGTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGTGATGCATGTATTTCTTCCCAGCTGTTCCGCACGTAAGCATTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATTAC
ATTACAGCCTGTATTCACTTGTCTAGAAGTTGTGAGAATTGACTGATTTGACTGTTGACATAATTGTAAT
GCATATATAACAATTGAAGCACTCCTTCTTCAGTTCTCAGCTCCTCTCATTTCAAGAAATATCCATT
TCAAGGTGCAGAACAGGAGTGAAAGAAAATAAGAAGAAAAATCCCTACATTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGTGGAAATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAACATCAACTTCATCAATTAGGAAGTATGGAACTAAGTTAAGGAAGTCCAGAAAGAACCAAG
ATATATCCTTATTTCATTTCAAAACTACTATGATAATGTGAAAGAAGATTCTGTTTGTGACCT
ATAATAATTATACAAACTTCATGCAATGTAATTGTCAGCAAAATTAAAGCAAAATTATTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAACACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTWHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMMSQRLESMVKNAYKSPLREEFVKSQVIKFSSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHPHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAPNPKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGCTTCCAGCTGAAGCCAATGCAGCCCCCGGCTCTCCGCGAAGAACGTTCCCTG
CCCCGATGAGCCCCCGCCGTGCCGCTCCCGACTATCCCCAGGGCGTGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTTGCCCTGGGAGTAGGATGTGGTGAAGGATGGGCTTCCTT
ACGGGGCTCACAATGGCAGAGAAGATTCCGTGAAGTGTCTGCCCTGCTCTAGCCCTCAA
TCTGCTTGGTTAATGTCCATCAGTGTGGCATGGTCTGGATGAGGAACGACTAA
ATAATGTTCTCACTTAACAGAGAACGAGGGTAGAGGAAGCAGTCATTGACTACTTCC
GTGGTTCATCGGTCTGGTGTGGCTTCCCTATCATTGTTGGGATGTTAGGATATTG
TGGAACGGTGAAGAAATCTGGCTTCTGCTGGATGGACTTGGAGTTGCTGTCAATTCT
GTGTAGAACCTGGCTGTGGCTTGGACATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGGCCAGGATGACAATTATGGATTACATGAGATATCGGTGGCTA
TGCTTGAATTTCAGAGAGGTTAAGTGTGGAGTAGTATTTCACTGACTGGTGG
AAATGACAGAGATGGACTGGCCCGAGATTCTGCTGTGGTAGAGAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTGTGGAGAAATGATTCC
TTTGAGAGGAACCAACAACGAGGTGCTGGGATCTCCATTGGGTGACACAAA
TCCTGGCATGATTCACCATTACTCTGCTCTGGCTCTGTATTATGATAGAACGGGAGGCTGG
ACAGACAAATGATGTCTGAAGAACATGACAACACTCTCAGCACCTGTCTGAGACT
GTTGAAACCAAGCCGTCAAGAACATCTTGAACACACATCCATGGCAAACAGCTTAAACACACT
TTGAGATGGAGGAGTTAAGAACATGTCACAGAAGAAACACAAACTTGTTTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTAGAAATATGTAGAAATAAAATGTTGCAAAAA
TAACACCTAACGATATAACTATTCTATGCTTAAATGAGGATGGAAAGTTGATGCAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTCATACCCACTGTGTA
TGTGTATGACTTTACTGAACACAGTTATGTTTGAGGCAGCATGGTTGATTGCAATTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGCCATAGTAAAGGGTGAATTACTTCTACCAA
CTAGTATATAAGACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTTTATT
CTCAGCGATCTATTCTCTGATGCTAAATAATTATATCAGAAAACCTTCATAATTGGTACT
ACCTAAATGTGATTGGTACTAAATATTCTACCACTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTGGTATATAAGTCTGTAAATCTGTATAATTCA
TTCACTGATAATGTTAGAACATAACCATATGAAAGGAAATTGTCCTGTATAGCATCATT
ATTTTAGCCTTCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTATTTAAATACCTAACACTAATTGGAAATTACCAACTGTGATACATAGGAATCATT
AGAATGTAGTGTGCTTTAGGAAGTTAATAAGAAAATTGCACATAACTTAGTTGATT
AAGGACTTGATGCTGTTCTCCCAAATGAAGACTCTTTGACACTAAACACTTTAAAAA
GCTTATCTTGCCTCTCCAAACAAGAACAGAACATAGTCTCAAGTCAATAAAATTCTACAGAAA
TAGTGTCTTTCTCCAGAAAATGCTTGAGAACATTTAAACATGTGACAATTAGGATT
CTTGTGTTATTCACTGATTAAATACTGTGGCAAATTACACAGATTAAATTGTTACAA
GAGTATAGTATATTATTGAAATGGAAAAGTCATTACTGTATTGTTGTTATTTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTCATAAAATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNVLTLAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLOQVLRLFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACATCCTCATCCTCTGATAAAGCCCTACCAGTGCT
GATAAAAGCTTTCTCGAGAGCCTAGAGGCCCTAAAAAAAAGTGCTGAAAGAGAAGGGGACAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGCAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGCCCTAGATCCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTGGATGAGTCAGGTCACCTCCCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCACCTCTGCATACAGAAAACCAGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCAGGGCATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGGAACACTGTCTCTGGAACACCAACGGTGGAAAGTCAGGCTGCCCTGGAGGCTACTATGTGATCGT
CTGACCAAGCCCAGCGTCTGCTTCACGTACTGTGGTCAATTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGACTGCACATGCCCTCAGGAACTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGAGTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGCAAGACTGTGAAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTTGCCTGGATCTGAGAAAGGCTACAGTGTGAAATGTCCCCGGGCTGGTGTG
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATGAAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCTGGAGCTTCTGACCAACACCTCCGTGCCAGGAGTGTCCAACGGCACCCATGTCACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCATGTGGTGAATGACAAGATTGTGCCAGCACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGAGCAGCGGGACTTCATCATCCGAACCAAGCAAGTGTGATCCCCGGT
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAAGGATACGTTCCAAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTCGTACTCCCTACTTGGCATTGAGCCCGTGGTGCACGTGAGCGG
TTGGAAAGCTTGGTGGAGAGCTGCTTGCCACCCCACTCCAAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGGTAAGCAGTACACATCCGGGATCACCTAGCAAAGCATTCCAGG
TCCCTGTCTCAAGTTGTGGCAAAGACCACAAGGAAGTGTGACTGCCGGTTCTGTGTGGAGTG
TTGGACCGCGTCCCGCTGTGCCACGGGCGATCCGCATCGACTGGAGGACTTAGTTCGTAGCCATACCTC
CGGTCTACAGGGCCAGACGCTAACAGCGGCCGATCCGCATCGACTGGAGGACTTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTGGAGCTCTCCCCCAACGCCCTCTAAGAACATCTGCCAACAGC
TGGGTTCAAGACTTCAACTGTGAGTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCATTCAAGTGGGCA
CAGGTACAGCAGCTGTAACAAATGTGGCCTGGTGGGTTCATCTTCTAGGGTTGAAAACCTAAACTGTCCA
CCCAGAAAGACACTACCCCACTTCCCTCATTTCTTCAACTTAAACACTCGTGTATGGTGCATCAGAC
CACAAAATCAGAAGCTGGTATAATATTCAGTACAAACCTAGAAAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTAAATATGTAATTAGTTACCTTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAAATTGGAAGTGTATCAATAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNNTVEVKACPGGYYVRLTKPSVCFH
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCShSCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRDLSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACCTGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATATTGT
GGTTGCTCTCATCAAGGTGATTCTGGATAAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGAACTGGTTCTCTGCCTGTTGACAACTTCACAGAACGCTCTCGTGGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACATCAAGTGGCCCTGTCTCAGGCTCCCTGGCTCCCTGCACGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGTGGGAGGAGGAGCCTCTGTGGATTCTTGGCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATTCCCTGGACCCCCACTGGTCCCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCACACTCACTTCTCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAACATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGTGTAACCAATCTGACCAAGTGGCATGTGGTGGCATCGTTAGCTGGCTATGG
CTGGGGGGGGCCCGAGCACCCCAGGAGTATACACCAAGGTCTAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAGTGTGGAGGCCCTCCTGCCCTGCCACCT
GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGGCCTCAATTCTGTAAGAGACCCCTCGCAGGCCAGGGCAGAGGAAGTCA
GCAGCCCTAGCTGGCCACACTTGGTGTCCCAGCATCCAGGGAGAGACACAGCCCACGTAAACAGGTCT
CAGGGGTATTGCTAACCCAAGGAACCTTCCCACACTACTGAATGAAAGCAGGCTGTCTGTAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGGCCCTGTCCGTCTCACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIALLSLASIIIVVVLIKVILDKYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEATACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVC CGSILDPHWLTAAHCFRKHTDVFNWKV RAGSDKL
GSFPSLA VAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGDGGPLMYQS
DQWHVVGIVSWG YGC GGPSTPGVYTKVSAYLNWIYNVWK AEL

Transmembrane domain:

amino acids 32-53 (type II)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTCACTGCAACTAAAAAGCCACAGGAGT
TGAACTGCTAGGATTCTGACTATGCTGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTTGTCTCTTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTACGTCT
GGAGGCAGTGACTCGGGCAGTCAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCCTGGCGTAGAAGGGATTTGACAAGCCCAGAAGATTCATAGGCATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGTCAACTACTGCCCTGGGCCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACCGCTGCCGCCATCTAACCTTCTATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTTGTGCCAATTATGGGTAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCTGCATTCACTGGACATGTGGGGAAAGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCGTCTGGAAAGGAGCAAGCAAAGTGACCAATTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCGCAAGACATAGCTGACAGG
GGCTAATGGCTAGTGGCCAGGAGGTCAAGCACAGGCTGCAAGGCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGAAATGCCCTCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTGTGATTTGAAACCCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGCTGTAATGCCAACAT
TTTGGGAGGCCAGGGCGGTAGATCACCTGAGGTCAAGACAGGCCCTGGCCAACATGG
TGAAACCCCTGTCTACTAAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAACACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLAELTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGCTCTCAGCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTTCCTGGCCCTAACTCTAATTGTCTT
GTTTGCCCCAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTACATGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAACGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAAAACGGATAACACTGG
CATCTACTCGTGGTCTCAAAAATGTTTATCAAAACTCAGATTAAGTGAATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTTCTTGAACAGTCAGT
ATTTGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCTGCCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCCTGCTGGATGAGAG
GTTATTGTTGATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGTCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGCTGTCTCATGCCCTG
TAACTGGTGGTGGCCCGATGCTGGGGAGGGCTTAATAGGAGGTTGAGCTAAATGCTAAAC
TGCTGGCAACATATAAAATGCATGCTATTCAATGAATTCTGCCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTCTCATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLIVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHF PANNEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTEENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTAGCTTCACACCTCGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAACAGTGGTGGCGTCCCTGTCATCCTGGGCT
GGCGGCTGCATCGGGCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAAGGCTTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCAATTGCTCAGGTCTTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTGA~~CTA~~ACTTGGATGTCCACAGCTAACATGTACACCGCATGGTGG
GATGGTGCAGACTGTCAGACCAGGTACACATTGGTGC~~GG~~CTGTTCTGGCTGGTCGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATGCCCTGCCGGGCTGGCACAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTCAAGGCCAGCAGTGGCTT
TGGGTCCAACACCAAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTATC
CTTCCAAGCAGACTATGTGTAATGCTCTAACAGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAACAAGGAGATCCCACATCTAGATTCTCTGCTTTGACTCACAGCTGGAAGTTAGAAAGCCT
CGATTTCATCTTGGAGAGGCCAATGGTCTTAGCCTCAGTCTGTCATAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTAGCTCACATTTCATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCCAGCTTACCCCCAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAGATTTGTAATCTCTCACATTGATGATCTCGGTTTCTT
ACACTGTGATCTAAAAGTTACCAAACCAAAAGTCATTTCTGAGGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCCACTGGAGCTCTTCTGT
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAATTAAAGTCCTAAATATAGTTAA
AATAAAATAATGTTAGAAAAATGATAACTATCTCTGTAATAGCTCACCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTGACATTGCTATATGGTACATTGTAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACTGTAATCTAGCAGTTGGAGGCTGAGGAGGAAGGATCACTTGAGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATACACCTGTAGTCCCAGCATTCCGGAGGCTGAGGTGGGAGGATCACTTGAGCCAGGGAGGT
TGGGGCTGCAGTGAGGCCATGATCACACCACACTGCACTCCAGCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTTAA

FIGURE 118

MSTTCQVVAFLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKA STGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTTG
GTGGTGGATGGGGCACAGTGGCTGTCAGTGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAACACTCTGGAGGACTGTGGATGAATTGCGTGGCAGGCTAA
CATCAGGATGCAGTCAGGATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGGCCAGAG
GACTGATGTGCTGCTCCGTGATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGTCAGGGCTGGAATCATCTTCAT
CATCACGGCATGGGGCTCATCCCTGTGAGCTGGGTGCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCGTTTTGGTGCACAGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCATCGCACACCAAAAAAGTTATCACACCGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTTAACTTAACTAAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTCAAATGGACCCCCAAGAAACTTTGATTACTGTTCTTAACGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTATGATTCTATAAGCTATTTCAGCAGAATGAGATA
TTAAACCAATGCTTGATTGTTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATTCTA
CTCTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAATGAAATGCCAGTCCATTACACTGAATAAAATAGAAACTCAACTATTGCTTTCAGGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTAAATTGTTAAAAACAGCTAGGGATAATGCTCCA
TTTATAATGAAGATTAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGATTAAATTAAACATTAAACGGCAGATATTGTCAGGGCTTGCATTCAAACGTCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGGTTAGGAAAGTG
AAAATTTTGTGTTGTATTGAAGAAGAATGATGCATTGACAAGAAATCATATATGTATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAAGAGCAGAAAAATA
TGTCTGGTTTCATTTGCTACCAAAAAACACAACAAAAAAAGTTGTCCTTGAGAAACTCACCT
GCTCCTATGTGGGTACCTGAGTCACAAATTGTCATTGTTCTGTGAAAATAATTCCCTCTGTA
CCATTCTGTTAGTTACTAAATCTGAAATACTGTATTGTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTATATAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCAGATTGATGCAGCCTGCAGCGGGCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCTCCGCCAGCTCCGCCTGCCAGCC
GGGAGCCCATGCGACCCCAGGGCCCCGCCGCCCTCCCGCAGCGGCTCCGCAGGCC
TGCTGCTGCAGCTGCCCGCGCCGTGAGCGCCTCTGAGATCCCCAAGGGAAAGCAAAAGGCGAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCAATGTATTCCGGTACACCTGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCGAGGGAAAGCTTGAGGGACTCTGGACACCCAAC
AACAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTCAAATAGTGTCTAAGAGTTGTTCACTGGCTCACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTCAGGACCTTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGGATTAGTGGATGTTGCTATCTGG
TTGGCAGTGTTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTGAGTTCTCGC
ATTATTGAAGAACTACCAAAATAATGCTTAATTTCACTTGTACCTCTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATAACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAAGTGTGATTACACTGTTAAATCTAGCATTATTCA
CTTCAATCAAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACTTCTTCAGTCACATT
CTCTCACCTATAATTGGAATATTGTTGGTCTTTGTTCTTAGTATAGCATT
AAAAAAATAAAAAGCTACCAATCTTGTACAATTGTAAATGTTAAGAATT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAA SPQR LRGL LLLLQL PAPSSASE I PKG KQKA QL RQ REVVDLY NGMCL QGPAG VPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDL GKIAECTFTK
MRSNSALRVL FSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAI IYLDQGSPEMNSTINI HRTS
SVEGLCEGIGAGLVDV AIVWVGTCSDYPKG DASTGWNSVSRIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGTACGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACTGCTGGCTATGGAATACAGATGTGCCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTAAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAAACTGTAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGATTTGCTGCGAATGCGGTGTTGGGATTATTGTTCTGGAG
TGTTCTGGTGGCTGGCAAAGAATAATGTTCAAATCGGTCCATCTCCAAGGGGTCCAATT
TCTTCTGGTGTAGCGAGGGCTGACTACAGTCAGCTGACAGGGCTGTCATGCAACTG
GCCCTAAAGCCAAAGCAGCTAACGGACCTTGAACAATAACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTATGCCCCACTGCTTACTGACAATG
CTTCTCTGCCGAACGAGGATGCCCTAACGGCTGAGCTGTAAGGCAAATGGTATATTGTA
ATCTCAGAAATTACAGGAGATACCCCTAACAGTATATCTGCTGGTGTAGGTTGCTTCGCT
ATAACAGCCTCAAAACTTAAGTATACTAACATTAAAGGCTCAACCAGCTCACCTGGTATAC
CTTGACCATAACATATCAGCAATATTGAGAAAATGCTTTAATGGAATACGAGACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCTATTCTAACATACCTTCAGACCTGTGACAATT
TACGGAACCTGGATCTGCTCTATAATCAGCTGATTCTGGGATCTGACAGCTTCGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGTCTAACCTGGAGAACCATCCCTGCGAATATTCA
AGACTGCCAACCTGGAACCTTGGACCTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGTGGCATGACTCACAAAGAACCTCACCTGGAGCACAATCAATTTCAGCTCAAC
CTGGCCCTTTCAAGGTTGGTCAGCCTTCAGAACCTTACTTGCAGTGGATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCCGAATCTGAGCAGCTCACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATACTCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGAGCAGAAATATTGCTCCCTGAAACTGGCTAAAAGTT
TTAAAGGTCAAGGGAGAATACAATTATCTGTCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAAGAACCTACAGCATCTGTCAGGAAACTACAGAGAGGTTGATCTGGCAG
GGCTCTCCAAAGCCGACGTTAACGCCAACGCTCCCAGGCCAGAACATGAGAGCAAACCCCTT
TGCCCCGACGGTGGAGCCACAGAGCCGGCAGAGACCGATGCTGACGCCAGCACATCT
TTCCATAAAATCATCGGGGAGCGTGGCCTTCTGTCGTGTCGTGTCATCTGCTGGTTAT
CTACGTGTCATGGAGCGGTACCCCTGCGAGCATGAAGCAGCTGCCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAGAAAAGACAGTCCCTAACGCAAATGACTCCCAGCACCCAGAACATTATGTA
GATTATAAACCCACCAACCGGAGACCAGCGAGATGCTGCAATGGGACGGGACCCCTGCAACCTA
TAACAAATGGGCTCCAGGGAGTGTGAGGTATGAAACCATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGGTGTCTTATTGAACTCTGGTACTATCAAGGGAACGCCGATGCCCCCCTCCCC
TTCCTCTCCCTCACTTGGTGGCAAGATCCTCTGTCGTTAGTGCATTCAATAACT
GGTCAATTCTCTCATACATAATCAACCCATTGAAATTAAATACCAATCAATGTGAAGCTT
GAACTCCGGTTAAATATAACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTCGCAATT
GTTTAAGATAAAACTCTTCTAGGAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNLSQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNALFPRLVSLONLYLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLFIGQEILDWSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTCGCTACTGCTGAATGTCCGTCCCGGAGGAGGGAGGGCTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGC^GCGCTACCGTGGCCGAGCT
AGCAACCTTCCCCTGGATCTCACAAAACTCGACTCCAATGCAAGGAGAACGAGCTTTGCTC
GGTGGGAGACGGTGAAGAGAACTGCCCCTAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGT^CGAATGGTCACATATGAACATCTCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTT
TTAGCCAATCCAACTGACCTAGTGAAGGTT^CAGATGCAAATGGAAGGAAAGGAAACTGGAAGG
AAAACCATTGCGATT^CGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCAATATAACAAAGAGCAGCACTGGTGAATATGGAGATTAA
ACCACTTATGATAACAGTGAAAACACTACTTGGTATTGAATACACCAC^TGGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAAATCATCG
ACTGACTGCTTGATT^CAGGCTGGTAAGGTGAAGGATT^CATGAGTCTATATAAAGGCTTTACC
ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAACTCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGRKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSrimnQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCAGGGCGGTGGGCCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCACCGCTGGCGCTGGTACGGGGGCTCGGGGGCATCGGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCTACCCGGACTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCCTGCTCTCAGGCAGCACAGTGGTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGAAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATGTCGCCACCGAGTGTAA
CCCCTGTCGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTGCCCTGCCTCTGGATTTAGGTGTTGAGTTCTGGAT
CACGGGATACCACTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATTTGTCAAATTGCTCAGTTGAAATGTGAAAAATGGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGTGCCCTGGGACTTGGCTTGTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTGTGACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTTATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGCCCTGGGACTGCACCCCTCTCCCCCTATCTATCTCCTCTGGCTCCCC
AGCCCAAGTCTGGCTTGTCCCCCTGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTGATGGTACATTTAAGAAAAAGAAAAATCGAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIIEELAAECKSAGYPGTLI
PYRCDLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMCKLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC ATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCCTCCTTGACTGCGGGCCGTTCAAGGTGCAAGAGTCTCAGTTGCCGGGAGC
ACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGAAAAGGCCATGGAACTTGGGTGAATCACCGATGCCATTAAAGAGGGTTTCTGCCA
GGATGGAATGTTAGGTCGTTCTGTCTGCGCTGTTCAATTCAAGTAGCCACGCCACCTGTGG
CCGTTGAGTGCTTGAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCATTTATTAA
TTAATTAACTGATAGTTGACATATTGGGGTACATGTGATAATTGGATACTGTATA
TATATAATGATCAAATCAGGGTAACGGGATATCCATCACATCAAACATTATTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCTCAGCTTACTGCAAC
CTCTGCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCCACAATGCCCAACTAATTGTATTAGTAGAGACGGGGTTTGCCTGTTGCCAG
CTGGCCTTGAACCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGGCCACCGTGCCTGGCTAAACATTATCTTGTGTTGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCACTGACTTCATCCCCACTCCTCTATCCTTCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGCTGGCTATTCACTAACATAATGACTCCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSSLRGPRPRIPLVSCQPV
KGHGTLGESPMMPFKRVFCQDGTVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGACTAAAATTGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCCTG
TCCATCTGTGTGCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACACCAAATAATAATGCTGGGATTCCCTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATAACCTATAACCACAAACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAGAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTCTGTCCCCTGAATCACCTAGCACAATTCCCTGGGTTT
GCCAGGACTATAGAAGAACTACCGCTGGATGATAATCGCATATCCACTATTCATCACCATCTTCAAG
GTCTCACTAGTCTAAAACGCCTGGTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTTCACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGAATTCCCTGACTGCTGCACAGTAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTTCAAGATAACCACATCAATCGGTGCCCTTAAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCTAAATAAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTATGTGCCAAGCCCCAGAAA
AGGTTGGTGGGATGGCTATTAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTGCAATACCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCAGTGAC
CAAACAGCCAGATATTAGAACCCCCAAGCTCAACTAGGATCAACAAACACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCAATCTCTGGAAACTTGCTCTACCTATG
ACTGTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCCGCATTGGATCTATAACAGAAACAATTG
AACAGGGGAAACGCACTGAGTACTTGGTCACAGCCCTGGAGGCTGATTCAACCTATAAAAGTATGCTGGT
CCATGGAAACAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAGAACCTTACAAAACCCCAATTAC
TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCATTCCTGCTTGTGTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTCTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAACACACCATATTCCCTCTAATGGAATGAATCTGT
ACAAAAACAAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSKLRLVLDGNLLNNHGLGDKVFNFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAPAQQWPAPVTQPD
IKNPKLTQDKQTGSPSRKTITITVKSVDSDTIHISWKLALPMTALRLSWLKGHS PAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLYLFDETPVCETETAPLRMYNPTTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCAGGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCCATGTTGCCAGGGGAGGTGGGAAGGAGGTGGAGGGCG
TGCAGAGGCAGTCTGGCCTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTGCTCTGCTTAAGCAGCAGTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGCTGGACACTGCTGAAAGTAAAGCCACCATTCAGACACTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGGAAGAGCAGCTAAAAGTGTCCGGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCTGGGATGCTGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTACCTAAGAGAGTCCAGCTGACCCCTCCAGCCGG
GTTTGAAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTGGTGTACCCACGTTGGG
CCCAGGACTCATTCTCAGAGGAGAGAAGTGTACGCTGAGGAGGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCACCAACTGCTCTTCTCCTGGGCCAGAATGAGGGATGCACACAGG
GACCACCCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGAACTTGAACCGC
AGAGCTGAGGCCATGGATACGCCAACCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAAATGGCGGCTCTCGACTTCTACAAGCTCCGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGAGCCTGATGCTGAAAGATGAAAGATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGGAGGTGAAGAGGCAGAAAAAACATTTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTCTGCCTCATCCTCCCAGTAGCTGGACTACAGGAGCGTGCACCATACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGCTCGAACCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCAGATGGCTGCTCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCCGCTGAGACGGACGGTCCATGCCAGCTGCCGGAGGAGGAACAGACCC
TTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTCAACTGCAACTGAACTGAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPMASLGLLLLLTLPPPLWSSSLPGLD
TAESKATIADLILSALERATVFQRLPEINLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRAEAIGYAYPTRDI FMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQOHFSRRVKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYLANFYIFLVETGFHHVAHAGLELLISRDPPSGSOSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTCCCGCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCGAGTGTGGTAGACGGAGAACGACGTGGTTCTTAAG
ACAGATGGGAGTTGTGGTTCATGATATACTCTGGATCTTATGAGTGGAAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTCTGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTTCTCCTTATTGATATTGTGCTCTGCCCTAAAGTGGTCAACACAAGTGTAC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCGTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACTATTATGATATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEHHVGFLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMMVPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCGATTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGAAAGGGGAGGCATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGCGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTAGACCTGAAATCTCTTCACGGGAGGCTGGCAGT
TTTCTTAACCTCTGTGGCTCCAGATTTCAGGCCAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTGCTGCGTTTATCTCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGGAAACATTGACATCAGAATCTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAAGATCAGCAGCAGCCTGCCAATTCTTCTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGAAACCTCAGGCAGCAGTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAAGAGGAGGATGACCCCA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTG
CTTCCTGCATGATTGTCTTATGCATCCCCATCTAATTGAGACCATCTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTTATTTGCTATT
ATGTATTATTTTACTTGGACATGAAACTTAAAAAAATTCAAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAAACCTTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCACTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGGAATAAGTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTTCTTGCATACCAAAAAAA
AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDdrvFKNYQTPDHYTLRKIS
SLANSFLTICKDLRLSHAHMTCHC GEEAMKKYSQILSHFEKLEPQAAVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCGCCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTCCGCTACCTGTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTCGGGGATTCTTCCCGCTCCGGTCCGGTCCGGTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGCGCCGAACCCCTCGGCTGGAGCCAGTCTCAACTGGACCACGCTGCC
ACCACCTCTCTTCACTAAAGTTGTTATTGTTCTGATAGATGCTTGAGAGATGATTTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACCTGTGAAAAGGGACATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCGGCTTTCGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGAAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGCTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAGCATTITGTGGAAATTATGATGGAACACCTCATTTTCTGTCAGATTACACAGAGGT
GGATAATAATGTCAAGGAGATTGGATAAAAGTATTAAAAAGAGGAGATTGGGACATATTAAATCC
TCCACTACCTGGGCTGGCACCATTTGCCACATTTCAGGGCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCACTGCGAAGGAGAGA
GACGCCCTAACCAATTGCTGGTTCTTGTGGTACCATGGCATGTCAGTGTGAAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTAAATCAGTTCTGCGTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCATGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAAGATCTGGGTTGAGCATTAAATGTCAAGAAAGATTGCAATGG
GAACCTGGATCAGACTGTACTTGGAGGAAAGCATTCAAGACTCTATTCAACCTGGCTCCAAGG
TTCTCAGGCACTACCTGGATGCTCTGAAAGACGCTGAGCTGTGCTCTGAGTCACAAGTGGCCAG
TTCTCACCTGCTCTGTCAGCGTCCCACAGGCACTGCAAGAACAGGCTGAGCTGGAAGTCCCA
CTGTCATCTGGGTTCTGCTCTTTATTGGTATCCTGGTCTTCTGGCGTTCACT
CATTGTGTGCACCTCAGCTGAAAGTTCTGCTACTCTGTGGCTCTGCTGGCTGGCGCAGGCT
GCCTTCGTTACCAAGACTCTGGTTGAAACACCTGGTGTGCAAGTGTGCTGGCAGTGCCCTGGAC
AGGGGCTCAGGAAGGACGTGGAGCAGCCTTATCCAGGCTCTGGGTGTCCGACACAGGT
TTCACATCTGTGCTGTCAGGTCACTGCCTCAGTTCTGGAAAGCTAGGTTCTGCACTGTTAC
CAAGGTGATTGTAAGAGACTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGCAGCTGAGGGAAAGAAGAGACAATCGGCCGG
CACTCAGGAGGGTCAAAAGGAGACTTGGTCGCACCAACTCATCCTGCCACCCCCAGAACATGCA
GCCTCATCAGGTCCAGATTCTTCAAGGGCGACGTTCTGTTGGAATTCTTAGTCCTGGCC
TCGGACACCTCATCGTTAGCTGGGAGTGGTGGTGGAGGGCAGTGAAGAAGAGGGGGATGGTCAC
ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCTGCACAGCCCTCATCCCTTGGCTTGAGCCGTCAAGGGCCCTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCACTCAGGGCACAGGCTCTGGAGGCCAGGATGATCTGT
CCACGCTTGCACTCGGGCCATCTGGCTCATGCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLLPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHLHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLOSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPPLISSLAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACA_{ACTGCCAGCACCTTAAGACCAC}TCAACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA_{ACTACATACGCCAGAGAT}CTTCTTGCA_{TAGCCTCATCCTTGAGCT}
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGCTCTAAAGGGAGTTTGCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAGAA_{ACTGATGAAGCT}
GGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTATAGGGCTCAGGTGGCTCCTGGA
ACATGCTGGAGTCGGCGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCCAT_{TGAACGCC}TTCCCTCGCTA
ATTTGA_{ACTAATTGTATAAAAACACCAACCTGCTCACT}

FIGURE 142

MLLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGTCTTCCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGGAGACAAGTTG
TCCCACAGGGCTGTCAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGCGCTGCTGCTGCTCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCAACAGACTGGG
CACATCTTGAATTCCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACCTGCTTCTTCACCACAGCACCAGGCCCTGGATGACTCAGTTCAAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAACCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACCGGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGGCCCAGGTGGCTGAGAACCGCCCCGGGAGCCTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGGCGGAACCAGGGGGAGTCCAGAACAGCCAGAACAAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGGACTCTGCTCCCTCAAGCATTCGAGGTTGGATCTGGCTATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCAGCAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGGCATGCACTCGCACTGCCAA
TGCGCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCCTGATAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACTCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATIC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTATGTCACTTATGTCTCACAGGGACCACATCCAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHKRINKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKHLHQO
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCTGGTCCAGAGTCTCATT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTGGCCTCCCT
CTGTCTTCTTCCCTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGTAGGTCACTGAGTCTTAGTTTATTTTGAATTT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTGAC
TCCC CGCGCCCCAACCCCTGCTTATCCCTGACCGTCGAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCGCCAACACCACCCCTGGCTTCTCCTGTTACTCCTCCTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGCTGTGACCCAAGCCAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTTAGTGTCCCATTCAAGTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTGAAACCTGCTAAAGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCCTAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGTTTACAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAAACAGTGAACAGTATCTAACACA
TTAACCTTGACAAATGGCTGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAGCAGAGAGAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTGAAAACCTGGATGAAATGATTGCTCTCAGACAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCCTAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAAATTGAATGGTGAAGAAAACATGACAAAAGGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTAAACACTTCAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTAAAACATAGCTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDFKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPFKVTPMAAIQDGLAKGENETVSNTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCTTAAGGTGAAGAGATCAGC
GTGGTC~~CCC~~CAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCTACAGACTAAGAGCCAGTGAACATCATGG
CCAGTGCCTGTCTAGTGGGGTGGGGCAGGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTCTACCCGGGGCATGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGTGGTTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGCAGACTCACCCAGCTCCGAGAATGGTGGCTGGAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGGGGCAGAGCCAGCTCGGTGAGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTTGGTCCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGCCTACTGCTGTGGAACCTTGAAACATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGGAGTGGTGGGAATATTCTGCT
TAATGGTAACTGACAAGTGTACCTGAGCCCCCGCAGGCCAACCATCCCCAGTTGAGCCTATA
GGGTCACTAGCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGCCCTTGGCCCAAGCCCCAACCCCTTCCCTTAATCTGCCACTGTCT
TGCTACCTTCTATCTCTCCCTCATCATCTTGTGTTGGCATGAGGAGGTGGTGTAGTCA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCAA
GATAACAATCAAATCCCAGATGCTGGTCTCATTCACATGAAAAGTGTCTCATGACATATTGAGA
AGACCTACTTACAAGTGGCATATAATTGCAATTATTTAATTAAAAGATACCTATTATATATT
TCTTTATAGAAAAAAGTCTGGAAGAGTTACTCAATTGTAGCAATGTCAGGGGGCAGT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTTATTTCTACAATGAAGATGA
ATTCCCTGTATAAAAATAAGAAAAGATAATTCTGAGGTAAGCAGAGCAGACATCATCTG
TTGTCCTCAGCCTCCACTTCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTCTGGTTGG
TTGTTAGTAGTGTAGCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGAGTTGT
GTGGCTGGAACTCTGGTAAGGAACCTAAAGAACAAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCTGGGATTCCAAGGCAATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGCCCCCTCAAATTCACATCCCTTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGCAAGATGTAGTTAGTAAAGACAAGGTCACTGTTGGATGAAGGTAGACCTAAATTCA
GAACGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCCGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAAGTTTGCAAGCCACAAGCTAAGAAAGACCCAAGGATTG
ATCAGAAAGCTTGGAAAGAGGCAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTCCAGCCTCTGAACGAAGAAAAGATAAAATTCCGGCTTTAA
GCCACCAAGGATAATTGGTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCTGTTAACATCTGTGTTGCTCTCCACATGTAACCTAAAGTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACCTGAGCCCTCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATGATGTTGTTTT
AAGTTGCTCAGTTGGTCTAACCTGTTATGCAGCAATAGATAAAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDAISLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCCCTGGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCAAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTTGAAGAAAGTGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTATATGCAGGAGGTGGTGCCTCTGGCCAGGCTCAGCAACAGGCTAACAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAAGCATAGATATTATTGATAACATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGTCTTTCCATAAAAAAGATTACTTCCAT
TCCTTAGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTTATTATCATTTATTAAATATGGATTATTAT
AGAAACATCATTGCTACTTGAGTGTAAAGGCTAATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKVSSFLMGTLATSCLLLALLVQGGAAAPISHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLLIGEKLFHGVSMERCYLMQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTCACTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCAAGC
TGCCAGGTTGGGGCTGGGGCCAAGTGGAGTGAGAAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCACAGGTGGTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGAGGTGGAGCACTGTGCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCAGTGAAGATGGACCCCTAACAGCAGGGCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTGAACCGCTCCCCCAGGACCTGTACCACGCCGT
TGCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTCTACAGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGCGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGGCCAGGTGTACA
ACCACTGCCATGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGCTGTGGAGCAG
CAGGATCCCGGGACAGGATGGGGCTTGGGGAAACCTGCACTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGGCCCGGAAGCTGGTGTCTGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCTGTCTCTCCCTTCCATCCCTGCTACCCCTG
GCCCAAGCACAGGCACATTCTAGATATTCCCCCTTGCTGGAGAAGAAAGAGCCCTGGTTTATT
TGTTTGTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGGAGCTGTTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTHTYSHWPSCCPKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACC GTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAACATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTGATGAATGTAAGCTGGGACTCCGGCAGATGCCAGCATCCGTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAAGACCCCTGGTGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGCTATTCAATTGGGCCATAATATTCTAACATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTACCTCACCAAGGTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAACCTGTGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTACAACCACTCCCTGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATCGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAAACAGGCGTCCCTTCCCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAACATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTT
TCTACCACCAACTACTGCCCTTAAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAACCAATTGCGAGAGTGGTCACTGGCTGCCACT
AGTGGCAGAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTAAAAGAAGGCA
GCAGACAAAGTCGTTCTCTTCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTCCCTGCCTTAACCTTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGTCACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACCTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCVPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCTGCATGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCGGCAGCTCGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCGGAGGAAGCACCAAGGCTGCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCACTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16O

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVMSMRNIESRSTSPWNYTWTWDPNRYPSEVVQAQCRLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGACAAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACAGGGCTGACTGGGGTGTCTGCCCTGGTCTGGGGGGGGCAGCAC
AGGGCCCTCAGGCCCTGGGTGCCACCTGGCACCTAGAAGATGCCCTGTGCCCTGGTCTTGCTGTCC
TGGCACTGGGCCGAAGGCCAGTGGTCTTCTCTGGAGAGGCTTGCTGGGGCTCAGGACGCTACC
CACTGCCTCCGGGCTCTCTGCCCTCTGGGACAGTGCACATACTGCCCTGCCCTGGGACAT
CGTGCCTGCTCCGGGCTCTGGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGAGCGACTGTGACCTGTCTGGTGTGGCTGTCCACTGGCGTGATGGGACTGG
GAAGAGGCTGAAGATGAGGAAAAGTTGGAGGAGCAGTGACTIONGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCTACCCACTGCCCTGCCGTCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTGTGCAAGTTGGTCAGTCTGTGGCTCTGGTATATGAC
TGCTTCAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCTTACTCAGCCAGGTACGAGAA
GGAACTAACACACACAGCAGCTGCCCTGCCCTGGCTAACGTGTCAAGCAGATGGTACA
ACGTGCATCTGGTCTGAATGTCTTGAGGAGCAGCAGTCTGCCCTCTCCCTGACTGGAATCAG
GTCCAGGGCCCCCAAACCCCCGGTGGCACAAAACCTGACTGGACCGCAGATATTACCTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAGGTGTGGCCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCTGAGAGCTGGCTGCTGGACGCCAGCTGCCCTGGCGAGAAGCGGACT
GTGCTGGCGGGCTCCGGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAAGACTGCAGCTGCAGGAGTGCTTGCTGGCTGACTCCCTGGGCTCTAAAGA
CGATGTGCTACTGTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCCTGGAACCCA
GTGGCTGACTTCACТАCCAGCAAAGCCTCCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCACTGAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGAGCCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGCCCTCGTGTGGCTGCCCTACTTTGCCG
CTGCGCTTCCCTCATCCCTCTAAAAAGGATACCGCAAAGGGTGGCTGAGGCTCTGGAA
CAGGACGTCCGCTCGGGCGGCCAGGGCGCGCTCTGCTCTACTCAGCCGATGA
CTCGGGTTTCGAGGCCCTGGTGGCGCCCTGGCGTCGCCCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGGCCCTGTAACCTGAGCGCGCAGGGGCCGCTGGCTTGGTTACGCGCAG
CGGCCAGACCCCTGCAGGAGGGCGCCGTTGGCTTGCTCTCCGGTGCAGGGCT
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGGGGGGGCGCACGGCCGACGACGCCCTCC
GCCCTCGCTCAGCTGCGTGCTGCCGACTTCTTGCAAGGGCGGGCGCCGAGCTACGTTGGG
GCCCTCGACAGGCTGCTCCACCCGGACGCCGTACCCGCCCTTCCGCACCGTGCCCTCTT
CACACTGCCCTCCAACTGCCAGACTTCTGGGGCCCTGCAAGCAGCCTGCCGCCCCCGCT
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCAGGGCCCTTCAGCCAGCCCTGGATA
CATCCCCGGGACTCCCGCCGGGACGCGGGGTGGGACCGGGGCGGGACCTGGGGCGGGGA
CGGGACTTTAAATAAAGGCAGACGCTGTTTCTAAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSCRWLSDIDLCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLFSQAYPTARCVLLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDHSVRTNICPFREDPRAHQNLWQAARLRLTLQSWLIDAPCSLPAEAALCWRAPIGGDPCQPLVPPLSWENVTVDKLEFPLLKHPNLCVQVNSSEKLQLQECWLADSLGPLKDVLLETRGPQDNRSLCALEPSGCTSLPSKAstraarlgeyLLQDLQSQCQLQLWDDDLGALWACPMDKYIHKRVALVWLACLLFAAALSLLKKDHAKGWLRLLQDVRSGAAARGRAALLLYSADDSGFERLGVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVLLFSPGAVALCSEWLQDGVSVPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVVFPLSQLPDFLGALQQPRAPRSGRILQERAQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites..

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCAGCAACTTGA
AAACATCCTGACGTGGACAGGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGGGACTGGGCAAGAAGGGCTGTCAGGGATCACCCGGAAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTACCGCT
GTCAGTGCGGAGGCCGTCAAGCCAAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTAGATGATTTGTCATCCTA
CCCCCAGCAGCAATCCGTGCAGGGATGGCACCCTGAACCTGAAAGACATCTCCATGACCTG
TTCTACCAACTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGAAAGCAGAGAGA
ATATGAGTTCTCGGCTGACCCCTGACACAGAGTTCTGGCACCATCATGATTTGGCTTCCCA
CCTGGGCCAAGGAGAGTGCCTTACATGTGCCAGTGAAAGACACTGCCAGACGGACATGGACC
TACTCTTCTCCGGAGCCCTCTGTTCTCCATGGGCTTCCCTGTCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCCTGCACCTCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCCTGCGCTCATCAGGAGCACGTCTGATCCCTGTCTTGAACCTCAGCGGCCAGC
AGTCTGGCCCAGCCGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCAGATCACCTACTTAGGGCAGCCAGACATCTCATCCTCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTCTAAGGTCAGCCTCCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGCATGGAAGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTCAAGCTCAGAAAGGCCACAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAATATTGACCCAGC
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCAGCTCCCTCCTCTCAGTCCAGATCGAGGGCACCCATGTC
CCTCCCTTGCAACCTCTCCGGTCCATGGTCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCTGTGTGTCAGGATGAAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTCAAGGGCTGGCCCTGACTGTGCAGTGGAGTCTG
AAGGGAAATGGGAAAGGCTTGGCTTCTCCCTGTCCCTACCCAGTGTCAACATCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGCCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGGGGGGAGCTCTGGGAGGAGCTGTGAGACAAGCGCGTCTGAGCCTG
CAAGGAGAAATGACAGTGCAGGAGGAAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCCTCAGGGAAATTGCGCTCTCCCTGCTCTGTGATTGTC
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCTG
GCCTGGAAAAAGAACCCAGGCTGGAAAAGAACCAAGAGAACGGCTGGCAGAACCCAGAACACTGC
ACTTCTGCCAAGGCCAGGGCAGCAGGACGGACTCTAGGGAGGGGTGTGCCCTGCAGCTCA
TTCCCGCCAGGGCAACTGCCGTGACGTTGACGATTCTGCTTCAATTCTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCAGGAGTTCA
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAACACAC
TGTACTGATGTCAACATTGCAAGCTGCTGCCCTGGGTCAGCCCATCTGGCTCAAATTCCAGC
CTCACCAACTCACAAGCTGTGTGACTTCAAAACAAATGAAATGAGTCCCCAGAACCTCGGTTCTC
ATCTGTAATGTGGGGATCATAACACACTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTCTTAATAGTGCCTGGTACATGGCGAGTGCCTAAACGGTAGCTATTAAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRIRTRKSCNLTVETGNLTEYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTTCIS
KVRSIQMIVHPTPTPIRAGDGHLRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHVLIPIVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDWSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSRAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDELQPTELDSLFRGLALTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGGTGCCACAACATGG
CTGCGGCGCCGGGCTGCTCTCTGGCTGTTGCTGGCTGGGGCGCTCTGGTGGTCCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTGAAGACTTCACGGGCCCTGATTGTCGTTGTGAATTTA
AAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAAGATTGATCAAGGTACTTCATAAAACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCGCTTGAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTGGGGCGTGA
ACTTGACCTGTGCCTGAGCCGAGGCATTCAAGAGCTGATTCAAGAGGATGGAGAAGGTGTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGGCCACCCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATAAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTGAGCAAACAAAATATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC GDEEC SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPPEESKKAEEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDEGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGCTCTGCCTCGCCTAGGACATACACGGGACCCCCTAACCTC
AGTCCCCCAAACCGCACCCTCGAAGTCTGAACTCCAGCCCGCACATCCACCGCGCGCACAGG
CGCGGCAGGCAGGCGCAGGTCGCCGAAGGGCATGCGCGCAGGGGTGCGGCAGCTGGCTCGGC
GGCGGGAGTAGGGCCGGCAAGGGAGGCTGCATATTCAAGAGTGCAGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCACGCAACACCTGCTGCGCCACGGCGCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGCGCCGCGCTGCTCGCGGCCACGGAGCTCTGCGCGCGTGGTCAGC
GGCCAAAAGGTGTTGCTACTTCAGGAGGCACGCCCTGGCTGTGAGAGTGAGGGAGGAGTCCCTCA
GTCCAGCCGAGTGAGCTTCAGGAGGCACGCCCTGGCTGTGAGAGTGAGGGAGGAGTCCCTCA
GCCTTGAGAATGAAGCAGAACAGAAAGTTAATAGAGACATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTCCTGATGGTGAATTCTGGATAGGGCTTGAGGAATGGAGATGGCAAACATCTGG
TGCTGCCCAGATCTTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCGCGGAAGTGAAAAGTGTGTTGATGTATACCAACCAACTGCCAATCTGGC
CTTGGGGGTCCCTACCTTACAGTGAATGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCCAGAGATTAATCCAACAGCCCTGTTAGAAAAGCCTTATCTTACAATCAACCA
GAGACACCCATCAGAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTGGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCAAACAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAAGTGGCATGAAAGTATAATAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACATAGCTTGGAAATGGCTGAAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTTTATATGCTATTATTCA
TTTAAAGAATATGCTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTGTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCTTTATTCTTACCTTCACTAAGTGTATCTAGTCATGTA
TGTATATTGTATTGAAATTACAGTGTGCAAAGTATTACCTTCACTAAGTGTGTTGATAAAA
ATGAACCTGTTCAATATTATTATGGCATCTCATTTCAATACATGCTTTGATTAAAG
AAACTTAACTACTGTGTCAACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTGT
TTTCTCGAAATAATTCACTTTCAGCTCTGCTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCAATTGAGTGTGATATAAACCTCCTCAAACATTACTAGAGGCAAGGAT
TGTCTAATTCAATTGTCAGACATGTCCTTATAATTATTAGCTTAAACAGATT
TTGTAATAATGTAACCTTGTATAAGGTGCTAAACACTAATGCACTGCAATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGTCCTCACCGTGCCTATATAATGAGAAGCAGCTCTGA
GGGTTCTGAAATCAATGTTGGCTCTCTGCTTGGCCACTAAACAAAGATGGTTGTTGGGGTTGGG
ATTGACACTGGAGGGCAGATAGTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTACAAAGAGGTATGTGGTTGAGACAGGTGAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACAGACATTAGGAAGGAAAGGAACATGCAAAATCGTGTGAAAATGGGTTGG
AACCCATCAGTGATCGCATATTCAATTGATGAGGGTTGCTGAGATAGAAAATGGTGGCTCTT
CTGCTTATCTCCTAGTTCTCAATGCTTACGCTTGTGCTTCTCAAGAGAAAAGTTGTAACCT
CTGGTCTTCATATGTCCTGTGCTCTTTAACCAAAATAAGAGTTCTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHAFCRVVSGQKVCFADFKHPCYK MAY FHELSSRVS FQE ARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVIPTIPLLILVAFGTCCFQMLHKSKRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217